JOURNAL OF APPLIED PHYSICS

JOURNAL OF APPLIED PHYSICS

Vol. 18, No. 10 OCTOBER, 1947

In This Issue

Page

Fringing Flux Corrections for Magnetic Focusing Devices NORMAN D. COGGESHALL 855

Quarter Wave Coupled Wave-Guide Filters WILBUR L. PRITCHARD 862

Computational Method Applicable to Microwave Networks R. H. DICKE 873

Optics of Three-Electrode Electron Guns S. G. ELLIS 879

On the One-Dimensional Theory of Steady Compressible Fluid Flow in Ducts with Friction and Heat Addition BRUCE L. HICKS, DONALD J. MONTGOMERY, AND ROBERT H. WASSERMAN 891

Crystalline Aggregation of Cobalt Powder J. T. MCCARTNEY AND R. B. ANDERSON 902

Surface Studies of Glass. Part I. Contact Angles L. A. SPITZE AND D. O. RICHARDS 904

Apparatus for the Observation of the Kerr Effect with Microsecond Electrical Pulses WILBUR KAYE AND RICHARD DEVANEY 912

On the Localized Afterglows Observed with Long Sparks in Various Gases J. D. CRAGGS, W. HOPWOOD, AND J. M. MEEK 919

Erratum: Conductivity of Metallic Surfaces at Microwave Frequencies E. MAXWELL 928

Letters to the Editor:

Quantitative Analysis of Mixed Powders with the Geiger-Counter X-Ray Spectrometer ZIGMOND W. WILCHINSKY 929

Conversion Factor for kX Units to Angström Units ELIZABETH ARMSTRONG WOOD 929

In Every Issue

New Books 930

New Booklets 930

Here and There 931

Calendar of Meetings 932

Index to Advertisers xxxviii

BOARD OF EDITORS

ELMER HUTCHISSON, *Editor*

Associate Editors

R. BOWLING BARNES
OTTO BEECK
W. F. BUSSE
W. JAMES LYONS
PHILIP MORRISON
PAUL E. SABINE
FREDERICK SEITZ
J. R. TOWNSEND
JOHN VON NEUMANN

Publication of the American Institute of Physics

ADMINISTRATIVE STAFF

HENRY A. BARTON, *Director*

CLEVELAND NORCROSS, *Assistant Director*

RUTH F. BRYANS, *Publication Manager*

THEODORE VORBURGER, *Business Manager*

JOHN T. TATE, *Adviser on Publications*

GOVERNING BOARD

GEORGE R. HARRISON, *Chairman*

WALLACE WATERFALL, *Secretary*

GEORGE B. PEGRAM, *Treasurer*

E. U. CONDON
A. J. DEMPSTER
W. F. FAIR, JR.
FLOYD A. FIRESTONE
HARVEY FLETCHER
R. C. GIBBS
K. S. GIBSON
PAUL E. KLOPSTEG
A. H. PFUND
JOHN T. TATE
M. W. WHITE
A. G. WORTHING

The JOURNAL OF APPLIED PHYSICS, published monthly at Prince and Lemon Sts., Lancaster, Pa., is devoted to physics in its role as the science basic to other natural sciences and to the arts and industries. Previous to 1937 this publication was known as PHYSICS. The JOURNAL OF APPLIED PHYSICS publishes editorials, news of physicists, and reviews, as well as technical papers of applied physics.

Manuscripts should be submitted to Elmer Hutchisson, Editor, Case Institute of Technology, Cleveland 6, Ohio.

Subscriptions and orders for back numbers should be addressed to Prince and Lemon Streets, Lancaster, Pa. or to the American Institute of Physics, 57 East 55 Street, New York 22, N. Y.

Subscription Price

	U.S. AND CANADA	ELSEWHERE
To members and associate members of the American Institute of Physics	$6.00	$7.00
To all others	8.00	9.00

Back Number Prices

Complete set: Vol. 1, July, 1931–Vol. 17, 1946—$124.00
Yearly back number rate: $10.00
Single copies: $1.00 each.

Changes of address and proofs should be addressed to the Publications Manager.

Advertising rates supplied on request. Orders, advertising copy, and cuts should be sent to the American Institute of Physics.

Entered as second class matter January 22, 1937, at the Post Office at Lancaster, Pennsylvania, under the Act of March 3, 1879.
Accepted for mailing at the special rate of postage provided for in the Act of February 28, 1925, authorized May 2, 1932.

Copyright 1947, by the American Institute of Physics.

Journal
of
Applied Physics

Volume 18, Number 10 October, 1947

Fringing Flux Corrections for Magnetic Focusing Devices

NORMAN D. COGGESHALL
Gulf Research and Development Company, Pittsburgh, Pennsylvania
(Received March 27, 1947)

When a charged particle approaches a region of uniform magnetic field its trajectory is affected by the fringing field. Compensation for the effect of the fringing field may in some cases be achieved by the use of virtual field boundaries which are outwardly displaced from the actual boundaries. A definition of the virtual boundary in terms of the effect of the fringing field upon the trajectories is formulated. A procedure for calculating the displacement d of the virtual from the actual field boundary is given using the methods previously developed for calculating ion trajectories in non-uniform magnetic fields. Using this procedure several values of the displacement for various magnet dimensions and conditions have been calculated. A method of altering the geometry of sector-type mass spectrometers to correct for the effects of the fringing fields is proposed. It is shown that the angular and spatial separations of the individual ion trajectories in an ion beam of small angular spread as it approaches a uniform field region through a fringing field are approximately the same as for the ideal case wherein the magnet field changes discontinuously at the boundary.

IT is well known that the magnetic field of a magnet does not discontinuously go from its maximum value to zero as the boundary of the region between the pole faces is passed. Rather the field falls off gradually in conformation with the laws of potential theory, and the result is an appreciable amount of fringing flux for some distance away from the pole-face boundaries.

In the sector-type mass spectrometer, schematically illustrated in Fig. 1, the ions proceed from the ion source through a magnetic-field-free region to the magnetic analyzer. The magnetic analyzer consists of a uniform magnetic field perpendicular to the plane of the figure and with the shape of a truncated sector. After passing through the magnetic analyzer the ions again pass through a magnetic-field-free region on their way to the ion collector. In the derivation[1] of the focusing properties of this system, it is assumed that the magnetic field region has sharp boundaries such as delineated in Fig. 1. Actually no magnet provides sharp boundaries to the magnetic field region, and in calculating the geometrical arrangement of ion source, magnetic analyzer, and ion collector, the effects of fringing magnetic field must be taken into account. One means of doing this is to use virtual boundaries of the sector-shaped field region which are displaced outwards from the actual boundaries. In this method it is assumed that fringing flux corrections can be effectively made by such a simple displacement of field boundaries. It is the purpose of this article to provide an analytical method for the determination of d, the displacement of the

[1] W. E. Stephens, Phys. Rev. **45**, 513 (1934).

virtual field boundaries from the actual field boundaries. This displacement is considered under various conditions and several numerical cases evaluated. It is a further purpose of this article to consider the effectiveness of such a procedure as the use of virtual field boundaries for fringing flux corrections and to provide an alternative procedure which is believed to be more effective.

If we consider the magnetic field on the median plane, i.e., the plane parallel to and midway between the pole faces, the flux on it may to a reasonable approximation be regarded as a function of only one variable, the perpendicular distance from the boundary of the uniform field region. Thus the methods previously developed[2] for the calculation of ion paths in non-uniform magnetic fields may be used to determine analytically the trajectories of the central ray and others between the ion source and their penetration into the uniform field region.

In the past there have been indicated approximate methods of correcting for stray fields,[3-5] but they do not allow the exact calculations of the trajectories as will be described below.

In reference 2 it was shown that the exact determination of ion paths could be readily achieved for such cases where the ions move in a plane through a non-uniform magnetic field which is a function of only one coordinate and is everywhere perpendicular to this coordinate and to the orthogonal coordinate in the same plane. This was shown to be true for both Cartesian and polar coordinate systems, the magnetic field being a function of the radial distance in the latter. For the case at hand we may use the distance from the uniform field region and on a line normal to its boundary as the x coordinate in a Cartesian system and apply these methods.

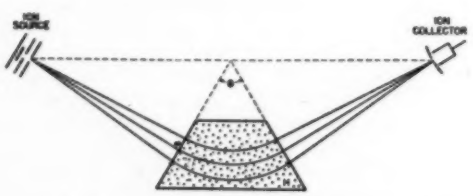


FIG. 1. Focusing scheme used in sector-type spectrometer.

In reference 2 it was shown that the ion trajectory is determined from the differential equation:

$$dy/dx = \pm f/(1-f^2)^{\frac{1}{2}}, \tag{1}$$

where $f = f(x)$ and is defined as:

$$f = (e/vmc)\int H(x)dx,$$

and v = velocity of ion in cm/sec., c = velocity of light, e = electrical charge of the ion in e.s.u., and m the mass of the ion in grams. For convenience in the present applications we define a function $h(x)$ as follows:

$$H(x) = H_0 h(x).$$

The function $h(x)$ then specifies the ratio of the magnetic field at any point to the maximum field H_0 in the uniform field region. With this we may rewrite (1) as:

$$dy/dx = \pm \bar{f}/(r^2-\bar{f}^2)^{\frac{1}{2}}, \tag{2}$$

where

$$f = \bar{f}(x) = \int h(x)dx \quad \text{and} \quad r = (vmc/H_0 e).$$

The quantity r is the radius of curvature of the ion path in the uniform field H_0.

It is to be noticed that the right-hand portion of Eq. (2) is a function of x only. Therefore, an integration, either numerical or analytical, may be made to determine y as a function of x and thus obtain the trajectory. If the function $h(x)$ is known only experimentally then $\bar{f}(x)$ may be determined by numerical integration and the resulting values used in a numerical integration of Eq. (2).

The application of the foregoing equations may be seen in Fig. 2. Here is plotted the trajectory of an ion as it passes from the ion source to and through a uniform magnetic field. Although the ion source is so oriented that its axis is perpendicular to the boundary of the magnetic field region, the fringing flux causes the central ray to actually enter the uniform field region with an angle other than 90° and at a

[2] N. D. Coggeshall and M. Muskat, Phys. Rev. 66, 187 (1944).
[3] A. J. Dempster, Phys. Rev. 11, 316 (1918).
[4] R. Herzog, Zeits. f. Physik 89, 447 (1934).
[5] J. Mattauch and R. Herzog, Zeits. f. Physik 89, 786 (1934).

point considerably displaced from the entrance point of the axis. The net effect of the fringing field is to so determine the trajectory that as the ion passes through the uniform field region the center of curvature is a distance d normal to the pole-face boundary. This is in contrast to the ideal case wherein the magnetic field would be discontinuous at the boundary and the center of curvature would be on the boundary. We may therefore define a virtual boundary as one parallel to the actual boundary and displaced from it so as to include as a point the center of curvature for that part of the trajectory in the uniform field region. We shall consider d at present only for the case wherein the axis of the ion source is perpendicular to the pole-face boundaries. As will be discussed in more detail later, d depends upon the distance of the ion source from the uniform field region and upon the angle of emergence of the ions from the source. In Fig. 2 the virtual boundary is represented by the dashed line which contains the center of curvature as a point. It is useful to know the displacement d, although as we shall see later these virtual boundaries do not allow the use of the ideal geometrical arrangements considered in reference 1.

The determination of the displacement d is readily made from Eq. (2). We note that $dy/dx = \infty$ for a value $\bar{x}$ such that $\bar{f}(\bar{x}) = \pm r$. Suppose the actual pole-face boundary is taken as the line $x=0$, then since in the x-coordinate system the center of curvature will be located r units from $\bar{x}$, it follows that:

$$d = (r - |\bar{x}|). \tag{3}$$

The absolute value is used here since $\bar{x}$ may have a negative value, depending on manner of using the coordinate system. To obtain $\bar{x}$ we must use the relationship

$$\int_{x_0}^{\bar{x}} h(x)dx = \pm r, \tag{4}$$

where x_0 is the value of the x coordinate at the ion source exit slit.

Thus the problem of determining d reduces to obtaining the function $h(x)$ and to carrying out the integration of Eq. (4). Actually for any particular magnet $h(x)$ should be determined experimentally. We may, however, get some idea of the order of magnitude of d and its change with variations of magnets by using the Schwarz-Christoffel transformation[6] to obtain $h(x)$. Figure 3 shows the type of magnet pole considered and the differential equation used for the conformal transformation. Here the magnet poles are considered to extend to ∞ in both directions perpendicular to the paper. They are in the form of jaws with the outer surface of each jaw being n times further from the median plane than the inner surface. For simplicity the transformation is made for the case shown by the cross-hatching in Fig. 3, i.e., one magnet pole as one surface of constant magnetic potential and the median plane as the other. The differential equation shown transforms the edge of the upper magnet pole onto the left-hand half of the real axis in the ω plane and transforms the edge of the other region of constant magnetic potential onto the right-hand half of the real axis. In the differential equation T is the usual orientation factor necessary for the transformation. The constants in the equation shown in Fig. 3, which expresses z in terms of ω, are given by:

$$a = (n + (n^2-1)^{\frac{1}{2}})/\pi,$$
$$b = (n - (n^2-1)^{\frac{1}{2}})/\pi,$$
$$c = -(((a+b)/2)\log(a-b) - (ab)^{\frac{1}{2}}\log(a-b)).$$

The transformation given in Fig. 3 allows the flux on the median plane to be calculated for various values of n. Changing the value of n

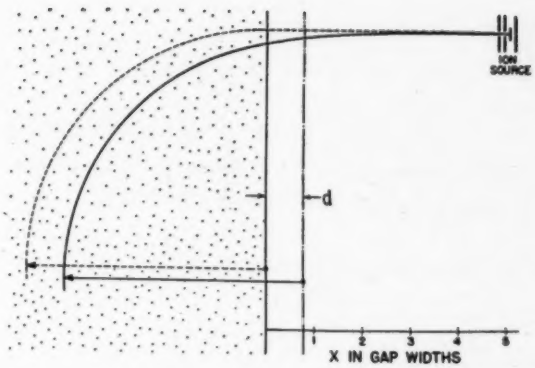


FIG. 2. Comparison between ion trajectories for the ideal case and for the actual case wherein the fringing flux is effective.

[6] See for example: W. R. Smythe, *Static and Dynamic Electricity* (McGraw-Hill Book Company Inc., New York, 1939), Chapter IV.

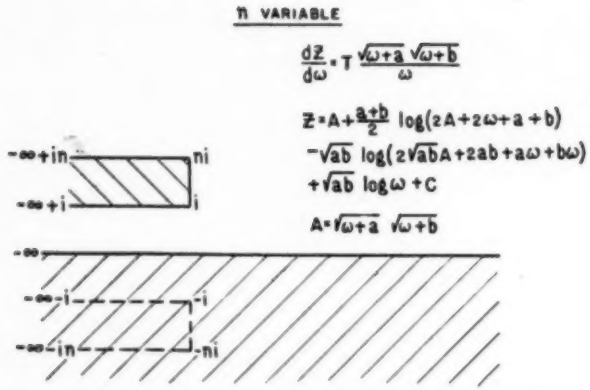


FIG. 3. Type of magnet pole construction considered and conformal transformation used.

corresponds to changing the thickness of the pole pieces relative to the gap width. The function $h(x)$ has been evaluated for several values of n, and the results are shown in Fig. 6.

In addition to the evaluation of $h(x)$ for several values of n the two extreme cases of $n=1$ and $n=\infty$ have been treated. Figure 4 shows the boundaries to be transformed and the differential equation for the extreme case of $n=1$. This case which is physically impossible corresponds to the pole pieces being thin sheets of magnetic material. The other extreme case of $n=\infty$, illustrated in Fig. 5, is also physically impossible but may be representative of magnets which have a large ratio of thickness to gap width. The results of the evaluation of $h(x)$ for these two extreme cases are also shown in Fig. 6.

It is somewhat surprising in inspecting Fig. 6 to see how little effect the variation of pole piece thickness has on the fringing flux. In this figure $h(x)$ is plotted against distance from pole-face boundary with the unit of distance as one gap width. Values of d obtained from using these evaluations of $h(x)$ cannot be expected to give true values of d for actual magnets. This is because an actual magnet will have finite dimensions in all directions and for it $h(x)$ will fall to zero more rapidly for large values of distance from pole-face boundary than shown in Fig. 6.

TABLE I. Variation of d with various values of n and r.

r/n	1	2	5	10	∞
5	0.69	0.84	0.92	0.95	0.98
10	0.81	0.97	1.07	1.14	1.20

If the dimensions are, however, reasonably large compared to the gap width, then the obtained values of $h(x)$ may be expected to be fairly accurate throughout the region wherein its greatest reduction in value occurs, i.e., up to about 2 gap widths.

In order to appraise the value of d obtainable for the two-dimensional cases discussed above the necessary integrations as in Eq. (4) were carried out. The results are given in Table I. Here d is given in gap widths as a function of the two variables n and r. In each case it applies to the central ray. Actually d is not directly dependent on r but rather on the distance of the ion source from the uniform field region. For the cases evaluated in Table I this distance was taken as equal to r, a situation which exists in a sector-type spectrometer when the sector angle is 90°. The values considered were: $r=5$ gap widths and $r=10$ gap widths.

As is evident from Table I, a large change in ion source distance will cause considerable change in d. Although the magnetic field strength is small at large distances it has a large cumulative effect. In determining $h(x)$ experimentally for a determination of d by numerical integration it is important to take the measurements under as near operating conditions as possible. For example, if an auxiliary magnet is used with the ion source to align the beam of ionizing electrons the experimental determination of $h(x)$ must include its effects.

It has sometimes been assumed that effective corrections for the influence of the fringing flux could be made by the simple use of the virtual field boundaries. Thus in locating the ion source and ion collector relative to the magnetic field the apex formed by the virtual field boundary is

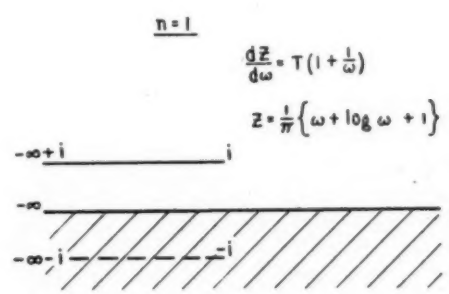


FIG. 4. Extreme case of magnet pole construction wherein $n=1$.

obtained by some approximate or graphical method. Such a method of correction assumes that the central ray will have its center of curvature located on the virtual sector apex.

We now wish to show that this method is not exact and that not only must the displacement of virtual field boundaries be taken into account but also the deviation of the trajectory from a straight line in the region between the ion source and the magnetic analyzer.

Figure 7 represents the geometry such as found in a 90° sector-type spectrometer with symmetrical location of ion source and collector. Here A represents the position of an ion source with its axis coincident with line AB. The effect of the fringing field is to cause the ions emerging from the source on the line AB to move away from it as they approach the magnetic analyzer. In the scale of dimensions used in constructing this figure AB is 5 gap widths which would be proper for a 90° instrument utilizing trajectories with radii of curvature also of 5 gap widths. The center of curvature for that portion of the central ray within the uniform field region is at D, the intersection of the virtual field boundaries DH and DI. Now if one were to use DH and DI as the new boundaries of an ideal magnetic field region, the ion source would be located at F which will be 5 gap widths from DH and which lies on the line perpendicular to ED, the bisector of the sector angle. However, it is clear that the location of the ion source at F will give a central ray with a center of curvature below and slightly to the right of D as the fringing field will still be operative. Thus the assumption that the use of virtual field boundaries would allow the use of

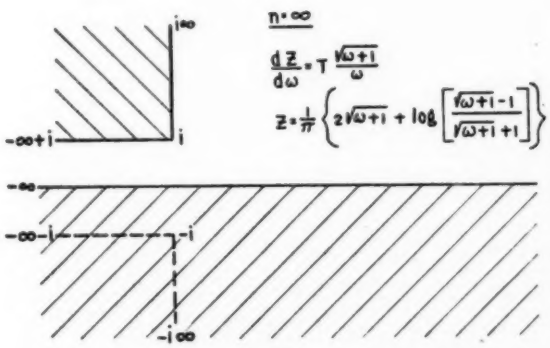


FIG. 5. Extreme case of magnet pole construction wherein $n = \infty$.

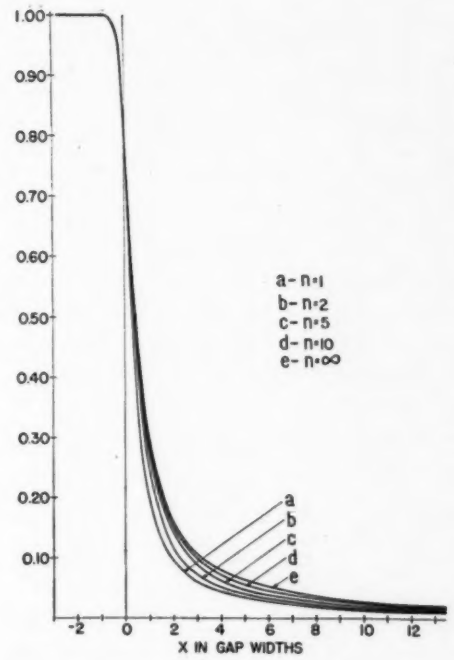


FIG. 6. Fringing magnetic field plotted as a function of distance from magnet pole faces for various values of n.

the exact geometrical relations prescribed for this type of instrument (reference 1) is seen not to be true. It is to be noticed that this result does not depend on the numerical value of d or of the sector angle, i.e., the use of virtual field boundaries and the exact geometrical relationships described above will always yield a center of curvature of the central ray offset from the virtual field boundary apex.

A method of correcting for fringing flux which depends upon an assumption concerning optimum focusing is now suggested. It is assumed that optimum focusing conditions are achieved when the ion source and ion collector are located the same distance from the actual field boundaries as would be used for the ideal case wherein there were no fringing flux and are so located relative to the sector apex that the central ray leaves the ion source in a direction normal to one pole-face boundary and enters the ion collector in a direction normal to the other pole-face boundary. Such a condition is illustrated in Fig. 7. There the distances AB and CG are equal to the radius of curvature for the uniform field region which is a condition necessary for the ideal 90° sector-type analyzer. One basis for the above

assumption about optimum focusing conditions is the close correspondence between the geometry stated and the geometry of the ideal case which was used in the derivation of the focusing properties of this type of analyzer. Another basis for the assumption is that as an ion beam of small angular spread reaches the actual pole-face boundary, after passing from the ion source through the fringing flux, the angular and spatial separations of the individual trajectories are approximately the same as would be obtained in the ideal case wherein there were no fringing field. Also the centers of curvature of the trajectories as they pass through the uniform field region will be separated by approximately the same distances as for the ideal case. These latter statements will be seen to be true later when we consider the changes in Eqs. (2) and (3) for ion trajectories of different angles of emergence from the source.

By way of demonstrating the above assumption concerning optimum fringing flux corrections let us suppose it is desired to make the corrections for a 90° sector-type instrument with symmetrical location of ion source and collector. Referring to Fig. 7, AB will be the distance of the ion source from the actual field boundary and will be equal to the radius of curvature chosen for operation. Supposing $h(x)$ to be available it will be possible to obtain $\bar{f}(x)$ and to use it in the integration of Eq. (2), which will give the deviation of the central ray trajectory from the line AB and will trace its entry into the uniform field region. When the trajectory reaches a point deep enough in the uniform field region that $h(x)=1$, the slope dy/dx may be evaluated and the location of point D, the center of curvature, determined. Now it is clear that we may expect the ion to leave the field in a trajectory exactly similar to the entry one if symmetrical conditions are provided. To do so a line which is at 45°, half the sector angle, with DH, the virtual field boundary, is projected to the actual field boundary. This locates E, the actual sector apex. Now it is seen that exactly symmetrical conditions are provided. The dimension of significance now is EB which locates the distances of the ion source and ion-collector axes from the actual sector apex. This is readily obtained in simple steps, using the location of point D relative to points A and E and to the point of entry of line AB into the uniform field region.

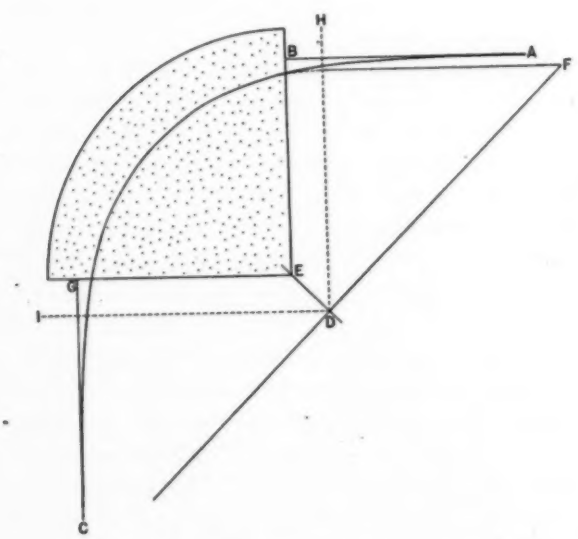


FIG. 7. Geometry of fringing flux corrections for sector-type instrument.

If it is desired to determine the flux corrections for an unsymmetrical system such that the angles θ and γ, as defined in reference 1, are unequal it may be done using the methods of above. In this case it would be necessary to determine the trajectories for two fringing field conditions, i.e., for the ion source and ion-collector sides, and match them in the uniform field region.

For the trajectories and geometry considered so far

$$f(x)=\int_{x_0}^{x} h(x)dx,$$

that is $\bar{f}(x_0)=0$ so that $(dy/dx)x_0=0$. This is the condition imposed on the central ray, i.e., that it leave the ion source normal to the slits (using such a case as in Fig. 7 where the ion source axis is parallel to the x axis and the pole-face boundary is parallel to the y axis). However, for ions emerging from the ion source at a small angle α with the normal to the slits we will have:

$$f(x)=\int_{x_0}^{x} h(x)dx+\bar{c},$$

where $\bar{c}$ is a constant to be correlated with the initial slope.[2]

From Eq. (1) we have then:

$$(dy/dx)x_0=\tan\alpha=\bar{c}/(r^2-\bar{c}^2)^{\frac{1}{2}}.$$

Since for small α (as would be true for an actual ion beam) $r\gg\bar{c}$ we obtain $\bar{c}=r\tan\alpha$.

The virtual boundary displacement for such a trajectory will be obtained from:

$$r=\pm\int_{x_0}^{\bar{x}}h(x)dx+r\tan\alpha.$$

Since $h(x)$ is equal to unity in the uniform field region we have for the virtual boundary displacement $d(\alpha)$ for a trajectory of starting angle α

$$d(\alpha)=d(0)+r\tan\alpha.$$

The separation then between the virtual boundary effective for a trajectory of starting angle α and the virtual boundary effective for the principal ray is seen to be $r\tan\alpha$ which is the same as is obtained for the ideal case. This result is part of the basis for the assumption above concerning optimum focusing conditions.

The separation in the y direction at the pole-face boundary between the central ray and one emerging from the ion source with angle α may be seen to be approximately the same as for the ideal case. To calculate this difference we integrate Eq. (2) for the two cases. We get then for the separation

$$\Delta y=\bar{c}\int_{x_0}^{0}\frac{dx}{(r^2-(\bar{f}+\bar{c})^2)^{\frac{1}{2}}}.$$

However, since $r\gg\bar{c}$ and since $\bar{f}(x)$ is always found to be small enough so that $(r^2-(\bar{f}+\bar{c})^2)^{\frac{1}{2}}=r$ to within about 2 percent or better for all of $0<x<x_0$ we have $\Delta y=\bar{c}x_0/r$ which is equal to $x_0\tan\alpha$, the same value as for the ideal case.

By use of Eq. (2) and the approximations above we may show that the angular spread of the trajectories as they enter the uniform magnetic field region is approximately the same for the actual case as for the ideal one. We shall denote the slope angle at $x=0$ as β. Then from Eq. (2) we have:

$$\tan\beta=(\bar{f}(0)+\bar{c})/(r^2-(\bar{f}(0)+\bar{c})^2)^{\frac{1}{2}}. \qquad (5)$$

By differentiation we get

$$(1+\tan^2\beta)\Delta\beta=r^2\Delta\bar{c}/(r^2-(\bar{f}(0)+\bar{c})^2)^{\frac{1}{2}}.$$

Using Eq. (5) and the approximation of above this reduces to $\Delta\beta=\Delta\bar{c}/r$. For the central ray and one emerging from the source at a small angle α we have $\Delta\bar{c}=\bar{c}=r\tan\alpha=r\alpha$ so that $\Delta\beta=\alpha$; i.e., the trajectories will have approximately the same angular divergence at $x=0$ as at the ion source. An actual calculation of the angular spread at $x=0$ for one of the cases for which d was evaluated above is illuminating. The conditions used were: $n=10$, $r=5$, and $\alpha=3°40'$. For this case $\bar{f}(x)$ was calculated for the region between the ion source and the actual field boundary using the values of $h(x)$ shown in Fig. 6. From this and from Eq. (5) it was determined that the angle at $x=0$ between the tangents of the two trajectories, was $3°44'$, i.e., the same as the starting angle to within less than 2 percent discrepancy.

The author is indebted to Dr. M. Muskat for several discussions of this problem and to Dr. P. D. Foote, Executive Vice-President of Gulf Research and Development Company, for permission to publish this material.

Quarter Wave Coupled Wave-Guide Filters*

WILBUR L. PRITCHARD

(Received March 24, 1947)

This paper considers the theory of band-pass wave-guide filters formed by cascading rectangular cavities separated by quarter guide wave-lengths. Formulas for resonant frequency, loaded Q, and insertion loss in the pass band are derived for a single stage using matrix theory and are in a form suitable for design purposes.

The general theory of "n" stages is worked out using matrix algebra, and it is shown that the insertion loss and phase shift of "n" stages are given by:

$$L_n = 1 + x^2 Un^2(x) \qquad \phi = \tan^{-1}[Tn(x)/Un(x)],$$

where $Tn(x)$ and $Un(x)$ are the first and rationalized second kind Tschebyscheff polynomials.

The specific design procedure is outlined including the design of a finite thickness inductive aperture.

I. INTRODUCTION

THE filters considered here are of the band-pass type, formed by rectangular cavities in cascade and separated from each other by quarter wave sections of line. The complete problem is divided into three sections. First, analysis of a single stage to find resonant frequency, phase shift, loaded Q, and loss in the pass band; second, study of the behavior of cascaded stages to find insertion loss out of the pass band, phase shift, and loss in the pass band for "n" stages; and third, the final design procedure.

The theory of a single stage is handled by use of matrix theory of four terminal networks[1] and the combined lumped element-transmission line equivalent circuit for the cavity. The cavity proper is represented by a length of ideal line, and the coupling apertures are represented by pure shunt susceptances. The latter representation is not strictly accurate, and a slight correction is necessary in the final design. The analysis of "n" stages in cascade is also handled by matrix theory and involves taking the "nth" power of a matrix.

Figure 1 shows the wave-guide configuration along with the equivalent circuit used in the analysis.

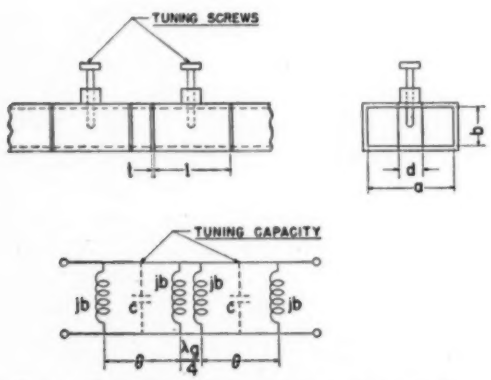


FIG. 1. Two stage filter and network equivalent.

II. PRELIMINARY NETWORK THEORY

A few brief relations based on the general circuit matrix of a quadripole are now derived. If the box in Fig. 2 represents any linear, passive network we can write:

$$E_1 = AE_2 + BI_2, \tag{1}$$

$$I_1 = CE_2 + DI_2. \tag{2}$$

The coefficients A, B, C, and D completely determine the network behavior. Evidently A and D are numerics, B is an impedance, and C is an admittance. Actually there are only three independent coefficients, since application of the reciprocity theorem to Eqs. (1) and (2) yields the relation.

$$AD - BC = 1. \tag{3}$$

Furthermore in the case of symmetrical networks we also have

$$A = D. \tag{4}$$

* The author wishes to acknowledge his indebtedness to R. M. Fano of M.I.T. for a preview of the manuscript of Vol. XXII of the Rad. Lab. Series which deals with the same problem from a different approach.

[1] L. A. Pipes, "Matrix theory of four terminal networks," Phil. Mag. **30**, 370–395 (November 1940); L. Brillouin, Revue Generale de L'Electricite **55**, 3–16 (January 4, 1946).

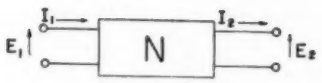


FIG. 2.

Since all the networks considered in the subsequent analysis are symmetrical, Eq. (4) will prove quite useful.

Rewriting Eqs. (1) and (2) in matrix form yields:

$$\begin{bmatrix} E_1 \\ I_1 \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} E_2 \\ I_2 \end{bmatrix}. \tag{5}$$

The matrix $\begin{bmatrix} A & B \\ C & D \end{bmatrix}$ is called the general circuit matrix and will henceforth be designated as $[u]$. The general circuit matrices for several basic network elements used in this paper are shown in Fig. 3.

In Fig. 4 we have a generator and load, normally matched to each other, separated by some four terminal network. The normalized generator and load impedances are assumed to be real and equal to one, since this is the condition commonly met in wave-guide problems. Since the impedances are all normalized, the following equations will not be dimensionally homogeneous.

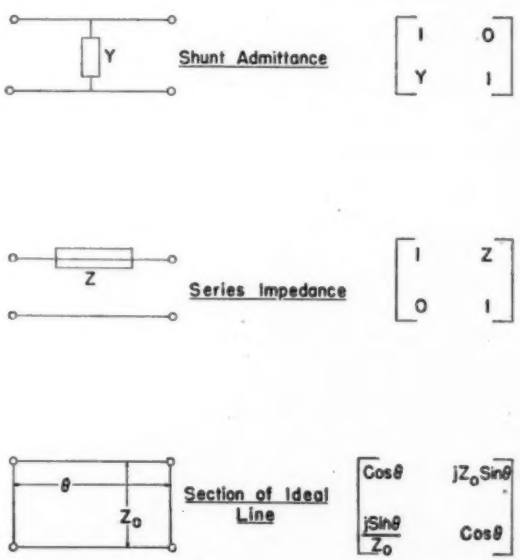


FIG. 3. General circuit matrices of several common networks.

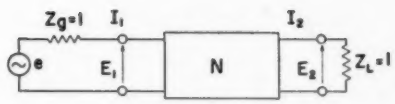


FIG. 4.

From Eqs. (1) and (2) we have:

$$Z_1 = E_1/I_1 = (AZ_L+B)/(CZ_L+D) = (A+B)/(C+D). \tag{6}$$

Using simple network theory we can find R, the ratio of the voltages across Z_L when directly connected to the generator and when connected through the network,

$$R = \tfrac{1}{2}(A+B+C+D). \tag{7}$$

R will be a complex quantity in general. The insertion loss, defined as the ratio of available power to the power delivered through the network, is found from the magnitude of R.

$$L = |R|^2. \tag{8}$$

The insertion phase shift, or phase shift in the voltage across the load resulting from the insertion of the network, is found from the real and imaginary components of R.

$$\tan\phi = [I(R)/R_e(R)]. \tag{9}$$

Equations (7)–(9) will enable us to determine the attenuation and phase characteristics of both single and multiple stage filters from the general circuit matrix.

III. ANALYSIS OF A SINGLE STAGE

Figure 5 shows the equivalent circuit for a single stage. We proceed with the analysis of a single stage by first determining the general circuit matrix. Since the circuit of Fig. 5 can be considered as a cascade combination of two shunt susceptances and a section of ideal line we find $[u]$ by matrix multiplication in the same order as the order of connection.

Referring to the table in Fig. 3

$$[u] = \begin{bmatrix} 1 & 0 \\ jb & 1 \end{bmatrix} \begin{bmatrix} \cos\theta & j\sin\theta \\ j\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} 1 & 0 \\ jb & 1 \end{bmatrix}. \tag{10}$$

Performing the indicated multiplication we have:

$$[u] = \begin{bmatrix} \cos\theta - b\sin\theta & j\sin\theta \\ j[(1-b^2)\sin\theta + 2b\cos\theta & \cos\theta - b\sin\theta \end{bmatrix}. \tag{11}$$

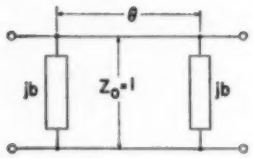


FIG. 5. Network equivalent of single stage filter.

From Eq. (7) we find R

$$R=(\cos\theta-b\sin\theta)+j[(1-b^2/2)\sin\theta+b\cos\theta]. \quad (12)$$

Taking the sum of the squares of the components of R and simplifying gives for the insertion loss:

$$L=1+b^2[\cos\theta-(b/2)\sin\theta]^2 \quad (13)$$

and for phase shift:

$$\tan\phi=\frac{[1-(b^2/2)]\sin\theta+b\cos\theta}{\cos\theta-b\sin\theta}. \quad (14)$$

Since the network is lossless a first definition for the resonant frequency is that frequency at which $L=1$. We have from Eq. (13)

$$1=1+b^2[\cos\theta-(b/2)\sin\theta]^2$$

Solving for θ

$$\theta=\arctan(2/b). \quad (15)$$

Equation (15) determines the resonant length of the filter in terms of the aperture susceptance. For large values of b which represent physically small apertures, θ is almost 180° or one-half wave-length. Equation (15) is plotted in Fig. 6, curve (2).

A second definition of resonance, and one which will be useful in the subsequent determination of the filter Q, is that frequency at which the input admittance is zero if the load admittance is zero. From Eq. (6) we have for the input admittance when $Y_L=0$

$$Y_1=C/A.$$

Substituting from the matrix in Eq. (11) yields

$$Y_1=j\frac{(1-b^2)\sin\theta+2b\cos\theta}{\cos\theta-b\sin\theta}.$$

Equating the numerator to zero we have:

$$(1-b^2)\sin\theta=-2b\cos\theta$$
$$\tan\theta=-2b/(1-b^2)$$
$$\theta=\arctan[2b/(b^2-1)]. \quad (16)$$

Equations (15) and (16) provide two different design relations for resonant frequency. For values of b corresponding to values of Q usually met in practice (treated in next section) the two formulas give the same result for $\tan\theta$ within normal slide-rule discrepancies.

To determine the physical length l of the cavity we use the relation

$$\theta=\beta l, \quad \text{where} \quad \beta=2\pi/\lambda_g. \quad (17)$$

β is the propagation constant of the guide and λ_g is the guide wave-length.

Thus

$$\tan 2\pi l/\lambda_g=2/b \text{ where } \lambda_g=\lambda/[1-(\lambda/\lambda_c)^2]^{\frac{1}{2}}. \quad (18)$$

In terms of ω, the angular velocity we have:

$$\tan\frac{\omega_0 l}{c}\left[1-\left(\frac{\omega_c}{\omega_0}\right)^2\right]^{\frac{1}{2}}=2/b, \quad (19)$$

where l = cavity length, b = aperture susceptance, λ = air wave-length, λ_c = cut-off wave-length, ω_0 = resonant angular velocity, ω = cut-off angular velocity, and c = velocity of light in vacuum.

Equation (19) is a design equation for determining the length of the resonant cavity. However, since there is an infinite number of pairs of values of b and l which satisfy Eq. (19) a second condition is required. This is the relation between the operating or loaded Q of the cavity and the aperture susceptance. To derive such a relation we find a lumped parallel resonant cir-

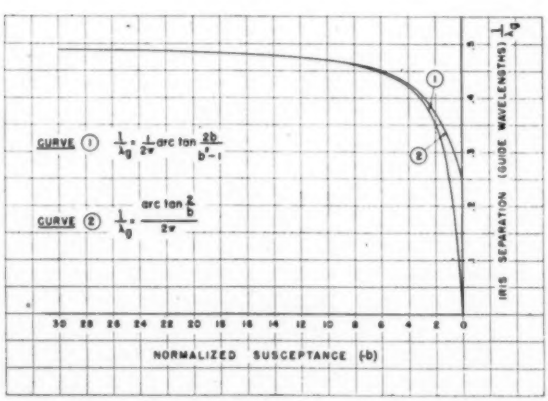


FIG. 6.

cuit which closely approximates the filter's behavior in the vicinity of resonance. In determining the filter's behavior near resonance we assume that the frequency consciousness of the apertures is negligible compared to the frequency consciousness of the length of line. In other words we assume constant b and consider variations in θ with frequency and their effects on the filter characteristics. Physically this assumption corresponds to neglecting the effect of the coupling holes in considering the characteristics *versus* frequency of a resonant cavity. Naturally this assumption is most valid for small coupling holes or large susceptances. Subsequently it will be seen that for usual values of Q the aperture susceptance will be high enough to make the above assumption quite justifiable.

Referring to Fig. 3 and the general circuit matrix we have for the input impedance, when the load impedance is infinite, from Eq. (1) and (2)

$$Z_1=(A+B/Z_L)/(C+D/Z_2)=A/C.$$

On an admittance basis we have from Eq. (11)

$$Y_1=C/A=j\frac{2b+(1-b^2)\tan\theta}{1-b\tan\theta}. \tag{20}$$

We assume $\theta=\theta_0+K\Delta f$, where θ_0 is the resonant wave-length and $K=d\theta/df$. By using the identity for the tangent of the sum of two angles and making the approximations $\tan K\Delta f=K\Delta f$ and $bK\Delta f<1$ Eq. (20) can be reduced to the form:

$$Y_1=jK\Delta f(b^2+1). \tag{21}$$

In Fig. 7 the input admittance with zero load admittance is the well-known relation:

$$Y_1=j(C/L)^{\frac{1}{2}}(\omega/\omega_0-\omega_0/\omega), \tag{22}$$

or for small frequency deviations $(\omega/\omega_0-\omega_0/\omega)=2\Delta\omega/\omega_0=\epsilon$,

$$Y_1=j(C/L)^{\frac{1}{2}}\epsilon. \tag{23}$$

Equating the two input admittances we have:

$$j(C/L)^{\frac{1}{2}}2\Delta f/f_0=jK\Delta f(b^2+1) \tag{24}$$

$$(C/L)^{\frac{1}{2}}=(Kf)(b^2+1)/2. \tag{25}$$

If we take the cavities as non-dissipative, then the Q for any given type of aperture depends only on the loading. Since there is a matched load $Z_0=1$ on each side of the cavity the shunt loading resistance in the equivalent resonant circuit is taken equal to $\frac{1}{2}$. Since

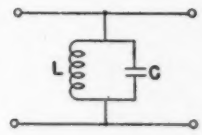


FIG. 7.

$$\begin{aligned} R_{\text{shunt}}&=Q(L/C)^{\frac{1}{2}}=\tfrac{1}{2}\\ Q&=\tfrac{1}{2}(C/L)^{\frac{1}{2}}\\ Q&=(Kf)(b^2+1)/4. \end{aligned} \tag{26}$$

To find K we take the first derivative of the function relating θ to f

$$\theta=\frac{2\pi lf}{c}[1-(fc/f)^2]^{\frac{1}{2}}. \tag{27}$$

After differentiating and simplifying

$$K=\frac{d\theta}{df}=\frac{2\pi l}{c}[1-(fc/f)^2]^{-\frac{1}{2}}. \tag{28}$$

Also, from Eq. (27)

$$f_0=\frac{c\theta_0}{2\pi l}[1-(fc/f)^2]^{-\frac{1}{2}}. \tag{29}$$

Substituting Eqs. (28) and (29) in Eq. (26) we find:

$$Q=(b^2+1)\theta_0/4[1-(fc/f)^2], \tag{30}$$

but

$$\theta_0=\tan^{-1}2b/(b^2-1).$$

Therefore:

$$Q=\frac{b^2+1}{4[1-(fc/f)^2]}\tan^{-1}\frac{2b}{b^2-1}. \tag{31}$$

For large values of b

$$\tan^{-1}\frac{2b}{b^2-1}=\pi \quad \text{and} \quad b^2\gg 1.$$

$$Q=\frac{b^2}{4[1-(fc/f)^2]}. \tag{32}$$

Equation (31) is plotted in Fig. 8. The calculation of Q using this expression neglects the effects of dissipative losses in the cavity. These

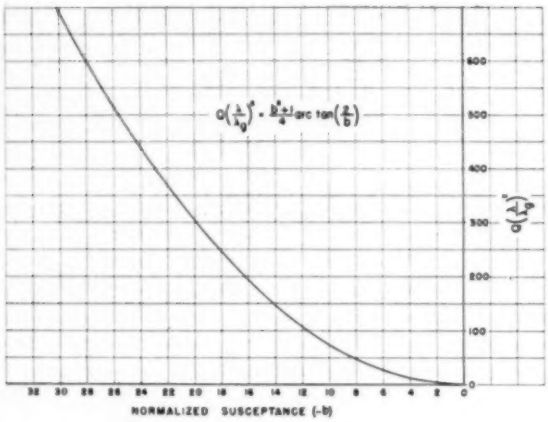


FIG. 8.

losses will have two chief effects on the filter performance. One will be to produce a slight lowering of the loaded Q according to

$$1/Q^1=1/Q+1/Q_0, \tag{33}$$

where Q^1 is the actual operating Q, and Q_0 is the unloaded Q of the rectangular cavity which can be calculated from ordinary resonant cavity theory.[2] The second effect will be to produce losses in the pass band. These losses can also be calculated from the unloaded Q in the following manner:

From the general circuit matrix for a shunt admittance we have:

$$R=(A+B+C+D)/2=(2+Y)/2$$
$$R=1+Y/2. \tag{34}$$

For a dissipative cavity Y is equal to the equivalent shunt conductance at resonance or:

$$Y=G_0=(1/Q_0)(L/C)^{\frac{1}{2}}. \tag{35}$$

But $(L/C)^{\frac{1}{2}}=\frac{1}{2}Q$, where Q is the loaded Q (Eq. (26)). Therefore

$$G_0=2Q/Q_0$$

and

$$R=1+Q/Q_0$$
$$L=|R|^2=(1+Q/Q_0)^2 \tag{36}$$

or

$$L_{\text{db}}=20\log(1+Q/Q_0). \tag{37}$$

Equations (33) and (37) can be used to correct for the effects of incidental dissipation in the cavity.

IV. QUARTER WAVE CASCADING OF FILTER SECTIONS

The simple filter of the previous section is only applicable where the desired response is approximately the same as that of a parallel resonant circuit. In many applications this is not sufficient, since the insertion ratio off resonance may not be high enough, or the width of the pass band may not be sufficiently great. In such cases it becomes necessary to cascade several filter sections. This cascading can be done in several ways. Either direct connection can be utilized, or the separate sections can be isolated by quarter wave sections. The successive sections can be tuned to the same frequency, or they can be stagger tuned. This report will consider only quarter wave coupled identical sections.[3]

We continue to use the lumped circuit equivalent for a single stage and consider the case of three quarter wave coupled stages (see Fig. 9).

The associated matrix is given by (refer to chart in Fig. 3):

$$[u]=\begin{bmatrix}1 & 0\\ Y & 1\end{bmatrix}\begin{bmatrix}0 & j\\ j & 0\end{bmatrix}\begin{bmatrix}1 & 0\\ Y & 1\end{bmatrix}\begin{bmatrix}0 & j\\ j & 0\end{bmatrix}\begin{bmatrix}1 & 0\\ Y & 1\end{bmatrix}, \tag{38}$$

Performing the indicated multiplication we have for $[u]$

$$[u]=\begin{bmatrix}1+Y^2 & Y\\ Y(2+Y^2) & 1+Y^2\end{bmatrix}. \tag{39}$$

In Fig. 10a we have a simple π section. To find its associated matrix we again multiply the matrices of its components.

$$[u]_\pi=\begin{bmatrix}(Z_p+Z_s)/Z_p & Z_s\\ (2Z_p+Z_s)/Z_p^2 & (Z_p+Z_s)/Z_p\end{bmatrix}. \tag{40}$$

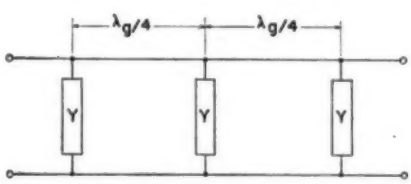


FIG. 9. Network equivalent of three cascaded stages.

[2] See, for instance, Sarbacher and Edson, *Hyper and Ultra High Frequency Engineering* (John Wiley and Sons, Inc., New York, 1943).

[3] Quarter wave coupled filters were first developed by R. M. Fano at Rad. Lab.

To find the equivalent π section for Fig. 9 we equate the elements of the two matrices given in Eqs. (39) and (40)

$$\begin{aligned} (Z_p+Z_s)/Z_p &= 1+Y^2 & Z_s &= Y \\ Z_s/Z_p &= Y^2 & Y_s &= 1/Y \\ Z_p &= Z_s/Y^2 = 1/Y \\ Y_p &= Y. \end{aligned}$$

Using the above relations we have a π section lumped circuit equivalent to the three section filter of Fig. 9. The value of Y is given by Eq. (23).

$$\begin{aligned} Y &= j(C/L)^{\frac{1}{2}}\epsilon = Y_p \\ Y_s &= 1/Y = -j(L/C)^{\frac{1}{2}}/\epsilon. \end{aligned} \tag{41}$$

The admittance of a lossless series resonant circuit is given by:

$$\begin{aligned} Y_{\text{series}} &= 1/j(L^1/C^1)^{\frac{1}{2}}\epsilon \\ &= -j(C^1/L^1)^{\frac{1}{2}}/\epsilon. \end{aligned} \tag{42}$$

By comparing Eq. (42) to Eq. (41) we see that Y_s can be replaced by a series resonant circuit where $C^1=L$ and $L^1=C$. Therefore we have for the final lumped equivalent the circuit of Fig. 10b.

This equivalent is evidently a band-pass filter with a center frequency given by $W_0=1/(LC)^{\frac{1}{2}}$ at which both series and shunt arms are resonant.

Similar lumped equivalents can be developed for any number of cascaded sections.

V. THEORY OF "n" SECTIONS IN CASCADE

The brief considerations of Section IV served to yield an equivalent lumped circuit for the particular case of three cascaded stages. The general case, to be considered now, involves "n" stages, and we propose to obtain an exact expression for insertion loss and phase shift. We proceed with this problem by considering the typical stage in a quarter wave coupled cascade. Fig. 11 shows the resonant circuit equivalent of the cavity with eighth wave lines on each side. Evidently a cascade of this type section will result in resonant circuits separated by quarter wave lines.

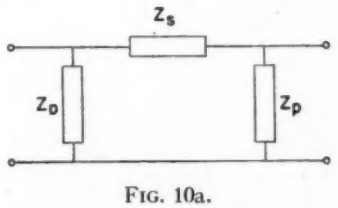


FIG. 10a.

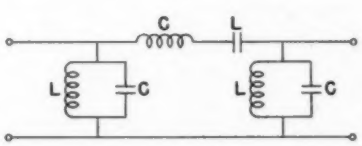


FIG. 10b. Lumped circuit equivalent of three cascaded stages.

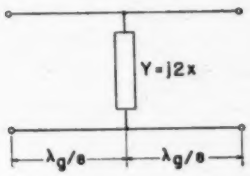


FIG. 11.

Again referring to Fig. 3 we have for the general circuit matrix:

$$[u]=\begin{bmatrix} \cos 45^\circ & j\sin 45^\circ \\ j\sin 45^\circ & \cos 45^\circ \end{bmatrix}\begin{bmatrix} 1 & 0 \\ Y & 1 \end{bmatrix} \times\begin{bmatrix} \cos 45^\circ & j\sin 45^\circ \\ j\sin 45^\circ & \cos 45^\circ \end{bmatrix}. \tag{43}$$

Substituting $Y=j2X$ where $X=Q\epsilon$, $\cos 45^\circ = \sin 45^\circ = (2^{\frac{1}{2}}/2)$, and performing the matrix multiplication yields

$$u=\begin{bmatrix} -X & j-jX \\ j+jX & -X \end{bmatrix}. \tag{44}$$

To raise the matrix $[u]$ to the nth power we must reduce it to the form

$$[u]=[K]\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}[K]^{-1}. \tag{45}$$

In this form $[u]^n$ becomes:

$$[u]^n=[K]\begin{bmatrix} \lambda_1^n & 0 \\ 0 & \lambda_2^n \end{bmatrix}[K]^{-1}. \tag{46}$$

λ_1 and λ_2 turn out to be the eigenvalues of the matrix $[u]$ and are found from the characteristic equation of $[u]$. The characteristic equation is formed by subtracting $[u]$ from the scalar matrix $[\lambda]$ and equating the principal determinant of the difference to zero. The roots of the

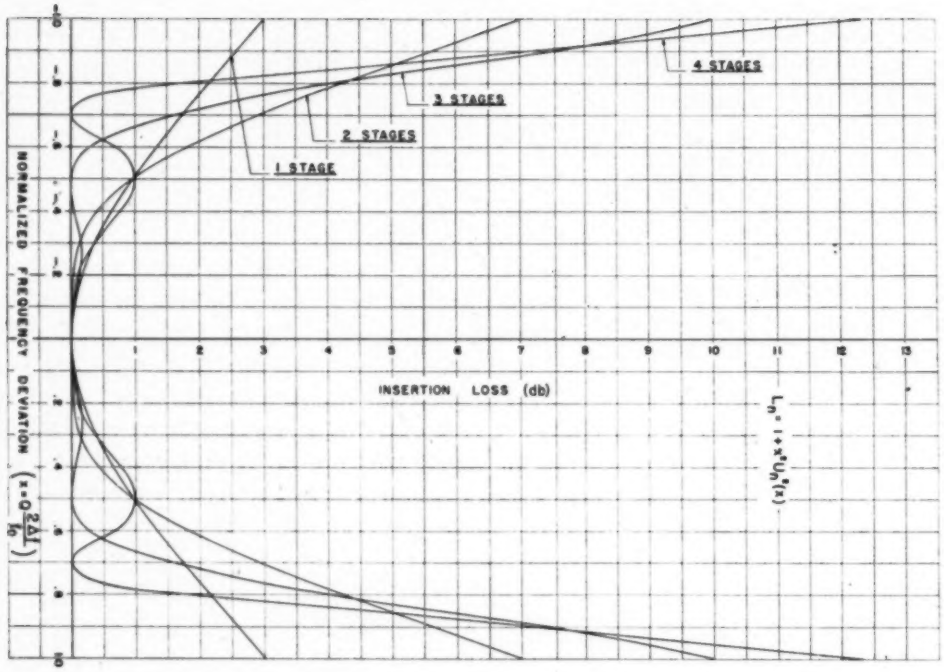


FIG. 12.

resultant quadratic (Eq. 47) then become:

$$\lambda^2+2x\lambda+1=0,$$
$$\lambda_1=-x+(x^2-1)^{\frac{1}{2}},\quad \lambda_2=-x-(x^2-1)^{\frac{1}{2}}, \tag{47}$$

Also:

$$\lambda_1+\lambda_2=-2x,\quad \lambda_1\lambda_2=1.$$

We make the substitutions $\lambda_1=\epsilon^a$ and $\lambda_2=\epsilon^{-a}$ which give:

$$\lambda_1+\lambda_2=2\cosh a,\quad a=\cosh^{-1}(-x). \tag{48}$$

The matrix $[u]$ can now be written in the form:[4]

$$[u]=\begin{bmatrix}\lambda_1+x & \lambda_2+x\\ j(1+x) & j(1+x)\end{bmatrix}\begin{bmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{bmatrix}\times\begin{bmatrix}\dfrac{1}{\lambda_1-\lambda_2} & -\dfrac{\lambda_2+x}{j(1+x)(\lambda_1-\lambda_2)}\\ -\dfrac{1}{\lambda_1-\lambda_2} & \dfrac{\lambda_1+x}{j(1+x)(\lambda_1-\lambda_2)}\end{bmatrix}. \tag{49}$$

The first and third matrices in the above product are inverses of each other, while the center matrix is a scalar matrix. That the product in the right-hand member of Eq. (49) is actually equal to $[u]$ can be verified by multiplication.

We now raise $[u]$ to the nth power by taking the nth power of the elements of the diagonal matrix, multiply the three matrices together and simplify to get:

$$[u]^n=\begin{bmatrix}\cosh an & -j\dfrac{\sinh a}{1+x}\sinh an\\ j\dfrac{1+x}{\sinh a}\sinh an & \cosh an\end{bmatrix}=\begin{bmatrix}A_n & B_n\\ C_n & D_n\end{bmatrix}. \tag{50}$$

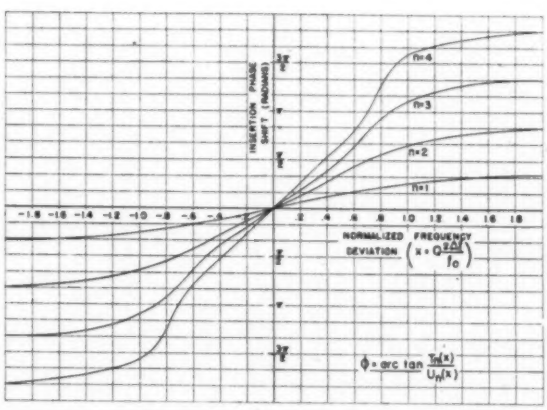


FIG. 13.

[4] The transformation and diagonalizing of a matrix carried out above are described in detail in Margenau and Murphy, *Mathematics of Physics and Chemistry* (D. Van Nostrand Company, Inc., New York 1943).

Equation (50) gives the general circuit matrix for "n" cascaded stages. Note that $A_n=D_n$ and $A_nD_n-B_nC_n=1$ which is to be expected. From $[u]^n$ we can now find the general expression for insertion less and phase shift: From Eq. (7) we have:

$$R_n=(A_n+B_n+C_n+D_n)/2.$$

Substituting from the matrix $[u]^n$ yields

$$R_n=\cosh an+j\frac{\sinh an}{\sinh a}. \tag{51}$$

For insertion loss we take the square of the magnitude:

$$L_n=|R_n|^2=\cosh^2 an+\sinh^2 an/\sinh^2 a. \tag{52}$$

Using the identity $\cosh^2 an-\sinh^2 an=1$ and substituting for Sinh$a$ we can write Eq. (52) in the form:

$$L_n=1+x^2\,\mathrm{Sinh}^2 an/(x^2-1) \tag{53}$$

or

$$L_n=1+x^2 \times[\mathrm{Sinh}\{n\ \mathrm{arc\ cosh}(-x)\}]^2/(x^2-1). \tag{54}$$

The expression in the bracket of Eq. (54) is equivalent to the rationalized Tschebyscheff polynomial of the second kind.[5] Rewriting Eq. (54) we have (see Appendix I):

$$L_n=1+x^2U_n^2(x). \tag{55}$$

Equation (56) gives the specific polynomials for $U_n(X)$ and L_n for values of "n" up to four along with a recurrence formula for U_n in case higher values of "n" are desired:

$$L_1=1+x^2, \qquad U_1=1,$$

$$L_2=1+4x^4, \qquad U_2=2x,$$

$$L_3=1+x^2-8x^4+16x^6, \qquad U_3=4x^2-1,$$

$$L_4=1+16x^4-64x^6+64x^8, \qquad U_4=8x^3-1,$$

$$U_{n+1}(x)=2xU_n(x)-U_{n-1}(x). \tag{56}$$

Figure 12 is a plot of Eq. (55) in decibels for $n=1, 2, 3, 4$. From the complex expression for R_n we can also determine the phase shift for

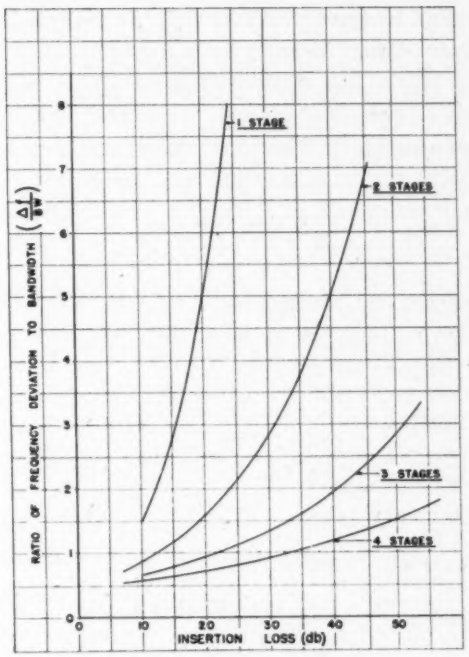


FIG. 14a.

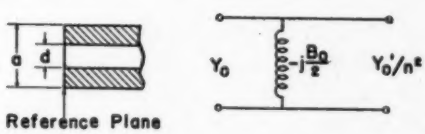


FIG. 14b. Network equivalent of inductive change in wave guide cross section.

"n" stages:

$$\phi_n=\tan^{-1}\frac{\sinh an/\sinh a}{\cosh an}$$

$$=\tan^{-1}\frac{\sinh[n\cosh^{-1}(-x)]/(x^2-1)^{\frac{1}{2}}}{\cosh[n\cosh^{-1}(-x)]}. \tag{57}$$

The expression in the denominator for Eq. (57) is a Tschebyscheff polynomial of the first kind, while the numerator is again a rationalized second kind polynomial. Rewriting Eq. (57) we have (see Appendix I):

$$\phi_n=\tan^{-1}\frac{-U_n(x)}{T_n(x)}. \tag{58}$$

Since the typical stage used in considering "n" stages cascaded had eight wave line sections on each end, when a particular number of stages is considered there is an extra eighth wave on

[5] E. Modelung, *Die Mathematischen Hilfsmittel des Physikers* (Dover Publications, New York, 1943).

each end of the composite filter. This total extra quarter wave introduces an added 90° phase shift. Since this added phase shift is not pertinent to the filter operation we subtract it from the expression in Eq. (58) yielding:

$$\phi_n = \tan^{-1}\frac{-T_n(x)}{U_n(x)}. \tag{59}$$

Specific expressions for ϕ_n and T_n for values of "n" up to four are given in Eq. (60) along with a recurrence formula for T_n

$$\begin{aligned}
&\phi_1 = \tan^{-1}x, && T_1 = x,\\
&\phi_2 = \tan^{-1}(2x^2-1)/2x, && T_2 = 2x^2-1,\\
&\phi_3 = \tan^{-1}(4x^3-3x)/(4x^2-1), && T_3 = 4x^3-3x,\\
&\phi_4 = \tan^{-1}(8x^4-8x^2+1)/(8x^3-4x), && \\
& && T_4 = 8x^4-8x^2+1,\\
& && T_{n+1}(x) = 2xT_n(x)-T_{n-1}(x).
\end{aligned} \tag{60}$$

Figure 13 is a plot of Eq. (59) for values of "n" up to four. In computing phase shift it is assumed that the phase shift at resonance is zero. Actually it will be some multiple of 90° because of the quarter wave sections and the lengths of guide composing the cavities. Also, if the individual stages are designed on an input match basis there will be a very slight extra phase shift at resonance. Generally this is negligible. Since one is usually only interested in phase shift variation about an arbitrary constant value, for f-m distortion considerations, etc., plotting ϕ_n about zero is quite useful.

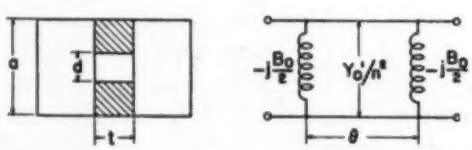


FIG. 15. Network equivalent of finite thickness inductive aperture.

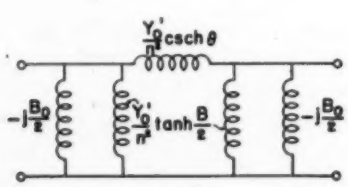


FIG. 16.

An approximate relation for the insertion loss of n stages, which will be accurate for values of x far off resonance, can be found from the Tschebyscheff polynomial by neglecting all but the highest power of x. This yields:

$$L_{n(\mathrm{db})} = 6(n-1)+20n\log x. \tag{60}$$

Equation (60) can be used to determine if a particular number of stages will have sufficient attenuation far out of the pass band.

VI. DESIGN PROCEDURE

In designing a filter for a particular application the first step is the choice of the proper number of stages. This is done with the aid of Fig. 14a. The ratio to band width of frequency deviation for a particular attenuation is plotted in Fig. 14 for $n=1, 2, 3, 4$. We define the band width of the filter as the difference between the frequencies of the half-power points. As an example, if an attenuation of 40 decibels is desired at a frequency two band widths removed from resonance, Fig. 14 shows that three stages are necessary. The curves in Fig. 14 can only be used to select the number of stages if band width is considered as distance between the half-power points.

After the number of stages has been determined the Q of a typical stage is found from Fig. 12. That is, from the normalized frequency deviation X which gives the desired band width, and from the band width and resonant frequency, Q is determined using the relation

$$X = Q[(f/f_0)-(f_0/f)] \tag{61}$$

or approximately

$$X = Q(2\Delta f/f_0). \tag{62}$$

The approximation in Eq. (62) is reasonably valid for values of frequency deviation as high as 10 percent which, at microwave frequencies, is much larger than normally considered.

Once Q is chosen and knowing λ/λ_g the aperture susceptance is found using Fig. 8, and then the aperture separation using Fig. 6. From the guide wave-length the actual mechanical cavity length is readily determined.

The last step in the design procedure is the determination of the actual aperture opening which will produce the desired susceptance.

Unfortunately this is a rather tedious process, since it is absolutely necessary to consider the finite thickness of the window.

The method is based on the equivalent circuit for a symmetrical change in cross section of a rectangular guide. This equivalent was derived by N. Marcuvitz at Radiation Lab.[6] Figure 14b shows the wave guide configuration, and the network equivalent referred to the indicated reference plane:

"a" is the cut-off controlling dimension of the guide, Y_0 the characteristic admittance of the first section and Y_0^1/n^2 the characteristic admittance of the second section of width d. B_0 is the susceptance of an infinitely thin aperture of opening "d" and can be found from Eq. (63) or from curves in the Wave Guide Handbook.

$$B_0/Y_0=\frac{\lambda_g}{a}\left[\cot^2\frac{\pi d}{2a}\right], \tag{63}$$

$$n^2=\frac{4\cos\pi d/2a}{\pi\,1-(d/a)^2}, \tag{64}$$

$$\lambda_g'=j2a/(1-(2d/\lambda)^2)^{\frac{1}{2}}, \tag{65}$$

$$Y_0'=Y_0[-j\lambda_g a/|\lambda_g'|d]. \tag{66}$$

$|\lambda_g^1|$ is the magnitude of the wave-length in the below cut-off section of guide of width d. Equation (65) for the imaginary wave-length in the cut-off guide is the usual expression for guide wave-length modified to take care of $(\lambda/\lambda_c)^2$ greater than one.

The above equivalent circuit can be extended to the case of a finite thickness aperture. Figure 15 shows the wave guide configuration and the equivalent circuit for symmetric inductive apertures.

In the above network the two lumped susceptances represent the abrupt changes in cross section at each end of the window, and the line section represents the short length of below cut-off line produced by the thickness of the window. This line section can be replaced by its exact π equivalent,[7] as shown in Fig. 16.

[6] *Wave Guide Handbook*, Supplement Rad. Lab. Rep. 41, Sec. 21a (January 23, 1945).

[7] Terman, *Radio Engineers Handbook*.

θ is given for a dissipationless line by Eq. (67)

$$\theta=j\beta l=\frac{2\pi}{|\lambda_g'|}t. \tag{67}$$

The network of Fig. 16 can be further simplified by combining the shunt susceptances of the π equivalent with $-jB_0/2$. The equivalent for the composite aperture is shown in Fig. (17).

$$B_a=B_0/2+|Y_0'/n^2|\tanh\pi t/|\lambda_g'|, \tag{68}$$

$$B_b=|Y_0'/n^2|\,\mathrm{csch}2\pi t/|\lambda_g'|. \tag{69}$$

Equations (68) and (69) completely determine the equivalent network for a finite thickness symmetric inductive aperture. To proceed with the design of the filter it is convenient to convert this π section to an equivalent T. This equivalent is shown in Fig. (18).

$$B_1=2B_b+B_a, \tag{70}$$

$$B_2=2B_a+B_a^2/B_b. \tag{71}$$

The shunt arm (B_2) of the T equivalent may now be used in determining the Q of the filter stage, and the series arm B_1 is used to make a slight correction in the resonant length of the filter. This correction is made by simply shortening the cavity an amount equal to the line length represented by the series susceptance B_1.

As a practical consideration it is generally desirable to design the filter at a frequency slightly higher than desired and then tune it down to the desired frequency with a capacitive tuning screw as shown in Fig. 1. This capacity will lower the resonant frequency as much as 5 percent without seriously changing the loaded Q.

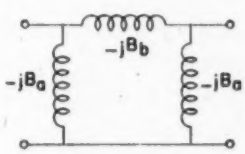


FIG. 17.

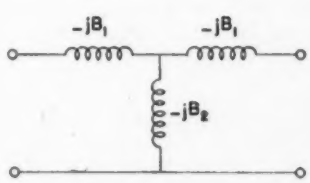


FIG. 18.

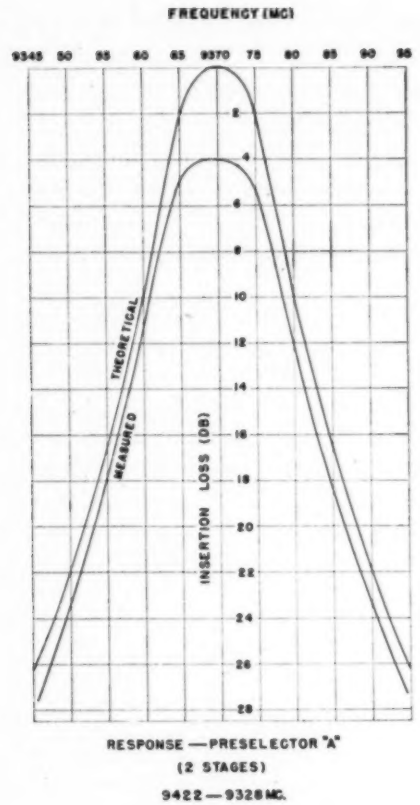


FIG. 19.

VII. EXPERIMENTAL RESULTS

Figure 19 shows a comparison between the calculated and measured characteristics *vs.* frequency of a two stage filter designed using the method outlined. The filter was designed for a frequency of 9422 Mc in $1'' \times \frac{1}{2}''$ wave guide and then tuned down to 9370 Mc with two capacitive probes mechanically ganged. The experimental curve is displaced downward from theoretical because of finite losses in the pass band, but its shape is substantially unaltered. It is important to note that the filter can easily be tuned 50 Mc below its natural frequency without materially altering its response.

Conclusions

A general theory for an "n" stage filter has been developed using matrix theory which, along with the design formulas for a single stage and the procedure outlined in the last section, can be used to design a composite filter.

APPENDIX I

We define the Tschebyscheff function of the first kind as follows:

$$T_n(x) = \cosh(n \cosh^{-1}x), \tag{1}$$

Letting $y = \cosh^{-1}x$ we have:

$$T_n(x) = \cosh ny \tag{2}$$
$$T_{n+1}(x) = \cosh(ny+y) = \cosh ny \cosh y + \sinh ny \sinh y \tag{3}$$
$$T_{n-1}(x) = \cosh(ny-y) = \cosh ny \cosh y - \sinh ny \sinh y. \tag{4}$$

Adding Eqs. (3) and (4) yields:

$$\begin{aligned} T_{n+1}(x) &= 2 \cosh ny \cosh y - T_{n-1}(x) \\ T_{n+1}(x) &= 2T_n(x)T_1(x) - T_{n-1}(x). \end{aligned} \tag{5}$$

Equation (5) is a recursion formula from which the Tschebyscheff polynomials can be formed, starting with $T_0(x) = 1$ and $T_1(x) = x$. If $-x$ is substituted for x then the sign of $T_n(x)$ changes for n odd and remains the same for n even. In Eq. (57) for phase shift the function in the denominator is actually $T_n(-x)$.

The second kind Tschebyscheff function is defined as follows:

$$U_n(x) = \frac{\sinh(n \cosh^{-1}x)}{(x^2-1)^{\frac{1}{2}}}. \tag{6}$$

The term $(x^2-1)^{\frac{1}{2}}$ is a rationalizing factor since $U_1(x) = \sinh(\cosh^{-1}x) = (x^2-1)^{\frac{1}{2}}$ and polynomials formed from a recursion formula will always have this radical as a multiplier. The recursion formula for $U_n(x)$, derived in exactly the same manner as the one for $T_n(x)$, is given by:

$$U_{n+1}(x) = 2xU_n(x) - U_{n-1}(x). \tag{7}$$

If $-x$ is substituted for x then the sign of $U_n(x)$ changes for n even and remains the same for n odd. In Eq. (54) giving the insertion loss for an "n" stage filter the rationalizing factor is already present in the denominator of the second term of the right-hand member, and since this term is squared the changes in sign of $U_n(-x)$ have no effect and we can write

$$U_n^2(x) = \left[\frac{\sinh(n \cosh^{-1}-x)}{(x^2-1)^{\frac{1}{2}}}\right]^2. \tag{8}$$

In Eq. (57) for phase shift the inverse tangent of ϕ is given by $(U_n(-x)/T_n(-x))$, which is equal to $-(U_n(x)/T_n(x))$ for n either odd or even. Since only relative phase shift is of interest this minus sign has been neglected in plotting ϕ.

A Computational Method Applicable to Microwave Networks

R. H. DICKE
Palmer Physical Laboratory, Princeton University, Princeton, New Jersey

(Received March 20, 1947)

A method is devised for computing the properties of a complex microwave network in terms of the properties of the circuit elements which in combination form the network. It is particularly suited to machine computation of the properties of circuits of such complexity that simpler, more direct methods fail. It is also applicable to low frequency networks. The elements of such networks may be regarded as interconnected by transmission lines of zero length. A numerical example is used to illustrate the method.

INTRODUCTION

IN the past the engineering of microwave circuits has consisted largely of a method of cut and try. It seems likely that microwave circuits of the future will consist of combinations of microwave elements each of whose properties are known. These microwave elements will probably be interconnected by transmission lines to form complex networks. Among the microwave elements used to synthesize a complex network one may mention tuners, line stretchers, tees, magic tees, directional couplers, and terminations. Figure 1 shows a block diagram of such a microwave circuit.

It is desired to calculate the properties of such a complex network in terms of the known properties of the various sub-elements which in combination form the network. In particular it is desired to calculate the amplitudes and phases of the waves leaving the network at each of the five feeding transmission lines in terms of the amplitudes and phases of the waves entering the network. For some purposes it is also necessary to know the amplitudes and phases of waves in the various interconnecting transmission lines.

As is well known, the complex amplitudes of waves leaving a junction of one or more transmission lines are linearly related to the amplitudes of waves entering the junction. This linear relation can be conveniently expressed by means of a matrix algebra formulation. See Eq. (1). The matrix algebra formulation will be used throughout the remainder of the discussion without going into a discussion of matrix algebra per se. (A concise development of matrix algebra can be found in Margenau and Murphy.[1])

[1] H. Margenau and G. P. Murphy, *The Mathematics of Physics and Chemistry* (D. Van Nostrand Company, Inc., New York, 1943).

$$\begin{pmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ \cdot \\ b_n \end{pmatrix} = \begin{pmatrix} S_{11} & \cdots & S_{1n} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ S_{n1} & \cdots & S_{nn} \end{pmatrix} \begin{pmatrix} a_1 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ a_n \end{pmatrix}. \quad (1)$$

Equation (1) is concerned with a junction of n-transmission lines. These lines are numbered consecutively from 1 through n. Reference planes are chosen in each of these lines. The complex number a_j represents the amplitude and phase of the wave incident on the junction in the jth line, the phase being measured at the reference plane. In a similar way b_j is a complex number representing a wave leaving the junction. The square matrix with elements S_{jk} is called the scattering matrix of the junction. Equation (1) expresses the fact that the b's are linearly related to the a's; Eq. (1) may be written formally as

$$\mathbf{b} = \mathbf{S}\mathbf{a}. \quad (2)$$

It should be noted that $\mathbf{a}$ and $\mathbf{b}$ in Eq. (2) are descriptions of the waves entering and leaving the junction, but that the matrix $\mathbf{S}$ is dependent only on the properties of the junction. The scattering matrix is simply related to the more familiar impedance and admittance matrices.[2]

DISCUSSION

The problem to be considered is the following: Given a set of microwave circuit elements, with scattering matrices $\mathbf{S}_1, \mathbf{S}_2, \cdots \mathbf{S}_n$, interconnected to form a complex network, what is the scattering matrix of the combined network? A

[2] Radiation Laboratory Technical Series, *The Principles of Microwave Circuits* (McGraw-Hill Book Company, Inc., New York, to be published), gives a discussion of the scattering matrix formalism and its application to microwave problems.

method will be presented for the calculation of the scattering matrix of a network composed of two interconnected junctions. It is clear that by iteration the method may be used to calculate a network with an arbitrary number of interconnected junctions.

Problem: To calculate the scattering matrix of two interconnected microwave elements.

Element No. 1 has scattering matrix $\mathbf{S}_1$ and is fed by p transmission lines.

Element No. 2 has scattering matrix $\mathbf{S}_2$ and is fed by q transmission lines.

Let there be r interconnections between the two elements. Number the transmission lines feeding the element (1) in such a way that the terminals of junction (1) number from 1 to p. Let the terminals of element (2) be numbered to run from $p-r+1$ to $p-r+q$. Let terminals in the range

$$p-r<n<p+1$$

be interconnected. For an example see Fig. 2. The scattering matrix for element (1) can be written

$$\mathbf{S}_1=\left[\begin{array}{c|c|c} & & 0 \\ \hline & & 0 \\ \hline 0 & 0 & 0 \end{array}\right]\begin{array}{l}\}p-r \text{ rows}\\ \}r \text{ rows.}\\ \}q-r \text{ rows}\end{array} \qquad (3)$$

$\underbrace{\quad}_{p-r}$ $\underbrace{\quad}_{r}$ $\underbrace{\quad}_{q-r}$

columns columns columns

Note that zeros have been added to make $\mathbf{S}_1$ a square matrix of order $p-r+q$. Equation (2) applied to junction (1) will be written as

$$\mathbf{S}_1\mathbf{a}_1=\mathbf{b}_1. \qquad (4)$$

In a similar way

$$\mathbf{S}_2\mathbf{a}_2=\mathbf{b}_2, \qquad (5)$$

where

$$\mathbf{S}_2=\left[\begin{array}{c|c|c} 0 & 0 & 0 \\ \hline 0 & & \\ \hline 0 & & \end{array}\right]\begin{array}{l}\}p-r\\ \}r\\ \}q-r\end{array} \qquad (6)$$

$\underbrace{\quad}_{p-r}$ $\underbrace{\quad}_{r}$ $\underbrace{\quad}_{q-r}$

Introduce an auxiliary matrix

$$\mathbf{A}=\left[\begin{array}{c|c|c} 0 & 0 & 0 \\ \hline 0 & \begin{matrix}1 & 0\\ & \ddots\\ 0 & 1\end{matrix} & 0 \\ \hline 0 & 0 & 0 \end{array}\right]. \qquad (7)$$

It is characterized by having zeros everywhere but along the diagonal of the central block. Also introduce

$$\mathbf{B}=\mathbf{I}-\mathbf{A} \qquad (8)$$

where $\mathbf{I}$ is the identity matrix.

The matrix $\mathbf{A}$ is useful in representing the connections between the two elements of the network. To see this, note that in Fig. 2 waves leaving junction (1) along transmission lines (4), (5) and (6) are also waves entering junction (2) and *vice versa*. This may be expressed formally in the following way:

$$\mathbf{A}\mathbf{a}_1=\mathbf{A}\mathbf{b}_2, \qquad (9)$$

$$\mathbf{A}\mathbf{b}_1=\mathbf{A}\mathbf{a}_2. \qquad (10)$$

From Eqs. (4), (5), (8)–(10) one obtains

$$\begin{aligned}\mathbf{b}_1&=\mathbf{S}_1\mathbf{a}_1\\ &=\mathbf{S}_1\mathbf{I}\mathbf{a}_1=\mathbf{S}_1\mathbf{A}\mathbf{a}_1+\mathbf{S}_1\mathbf{B}\mathbf{a}_1\\ &=\mathbf{S}_1\mathbf{A}\mathbf{S}_2\mathbf{a}_2+\mathbf{S}_1\mathbf{B}\mathbf{a}_1\\ &=\mathbf{S}_1\mathbf{A}\mathbf{S}_2\mathbf{A}\mathbf{a}_2+\mathbf{S}_1\mathbf{A}\mathbf{S}_2\mathbf{B}\mathbf{a}_2+\mathbf{S}_1\mathbf{B}\mathbf{a}_1. \qquad (11)\end{aligned}$$

Let

$$\mathbf{M}=\mathbf{S}_1\mathbf{A}\mathbf{S}_2\mathbf{A}, \quad \mathbf{N}=\mathbf{S}_1\mathbf{A}\mathbf{S}_2. \qquad (12)$$

Then

$$\mathbf{b}_1=\mathbf{M}\mathbf{a}_2+\mathbf{N}(\mathbf{B}\mathbf{a}_2)+\mathbf{S}_1(\mathbf{B}\mathbf{a}_1). \qquad (13)$$

It should be noted that the last two terms on the

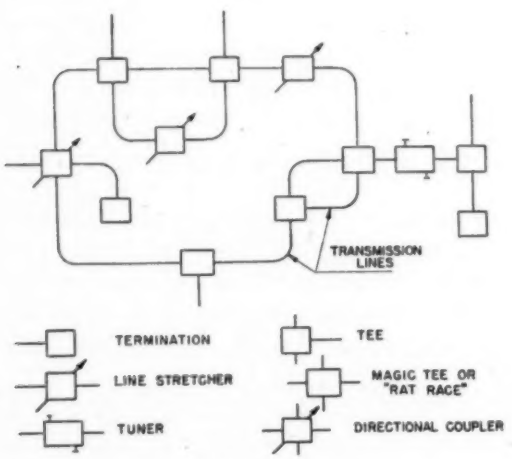


FIG. 1.

right side of (13) contain $\mathbf{Ba}_2$ and $\mathbf{Ba}_1$ as column vectors. From (7) and (8) it may be seen that $\mathbf{B}$ has all zeros in the central block. Hence if $\mathbf{B}$ operates on a column vector it replaces the r rows whose indices run from $p-r+1$ through p by zeros. Thus $\mathbf{B}$ has the effect of removing from $\mathbf{a}_1$ and $\mathbf{a}_2$ those elements which are unknown, namely, the wave amplitudes in the interconnecting transmission lines. The remaining elements of $\mathbf{a}_1$ and $\mathbf{a}_2$ are the incident waves in the transmission lines feeding the combined network and are assumed known.

The procedure which leads from (4) to (13) may be iterated to give the following sequence of equations:

$$\begin{aligned}
\mathbf{b}_1&=\mathbf{M}\mathbf{a}_2+\mathbf{N}(\mathbf{Ba}_2)+\mathbf{S}_1(\mathbf{Ba}_1),\\
\mathbf{b}_1&=\mathbf{M}^2\mathbf{a}_2+\mathbf{MN}(\mathbf{Ba}_2)+\mathbf{MS}_1(\mathbf{Ba}_1)\\
&\qquad+\mathbf{N}(\mathbf{Ba}_2)+\mathbf{S}_1(\mathbf{Ba}_1),\\
\mathbf{b}_1&=\mathbf{M}^3\mathbf{a}_2+\mathbf{M}^2\mathbf{N}(\mathbf{Ba}_2)+\mathbf{M}^2\mathbf{S}_1(\mathbf{Ba}_1)\\
&\quad+\mathbf{MN}(\mathbf{Ba}_2)+\mathbf{MS}_1(\mathbf{Ba}_1)\\
&\qquad+\mathbf{N}(\mathbf{Ba}_2)+\mathbf{S}_1(\mathbf{Ba}_1),\qquad(14)\\
&\ \ \vdots\qquad\qquad\vdots\qquad\qquad\vdots
\end{aligned}$$

Note that in this sequence of equations, with the exception of the 1st term, all terms on the right side have column vectors operated on by $\mathbf{B}$.

The Eq. (14) may be rearranged as follows:

$$\begin{aligned}
\mathbf{b}_1&=\mathbf{M}\mathbf{a}_2+[\mathbf{N}(\mathbf{Ba}_2)+\mathbf{S}_1(\mathbf{Ba}_1)],\\
\mathbf{b}_1&=\mathbf{M}^2\mathbf{a}_2+(\mathbf{M}+\mathbf{I})[\mathbf{N}(\mathbf{Ba}_2)+\mathbf{S}_1(\mathbf{Ba}_1)],\qquad(15)\\
\mathbf{b}_1&=\mathbf{M}^3\mathbf{a}_2+(\mathbf{M}^2+\mathbf{M}+\mathbf{I})[\mathbf{N}(\mathbf{Ba}_2)+\mathbf{S}_1(\mathbf{Ba}_1)].\\
&\ \ \vdots\qquad\qquad\vdots\qquad\qquad\vdots
\end{aligned}$$

Let the characteristic equation of the matrix $\mathbf{M}$ be

$$\sum_n C_n\lambda^n=0. \qquad (16)$$

It will be shown in the appendix that the term C_0 vanishes in (16) and furthermore that there can be no more than $r+1$ non-vanishing terms.

By Cayley-Hamilton's theorem the matrix $\mathbf{M}$ satisfies its characteristic equation and

$$\sum_n C_n\mathbf{M}^n=0. \qquad (17)$$

If possible, it is convenient to normalize the coefficients in (16) in such a way that

$$\sum C_n=1. \qquad (18)$$

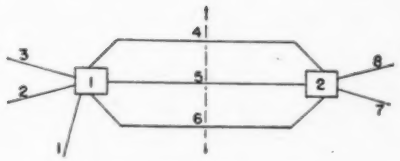


FIG. 2. $p=6$, $q=5$, and $r=3$.

If the sum of the coefficients of the characteristic equation vanishes, this normalization is impossible. This case will be discussed in the appendix.

Multiply Eqs. (15) by the coefficients C_n and sum over the index n.

$$\sum_{n=1}^{r+1} C_n\mathbf{b}_1=\left[\sum_{n=1}^{r+1} C_n\mathbf{M}^n\right]\mathbf{a}_2+\sum_{n=1}^{r+1} C_n\mathbf{S}_1(\mathbf{Ba}_1)$$

$$+\sum_{n=1}^{r+1} C_n\mathbf{N}(\mathbf{Ba}_2)+\sum_{n=2}^{r+1} C_n\mathbf{MN}(\mathbf{Ba}_2)+\cdots. \qquad (19)$$

Substituting (17) and (18) in (19)

$$\mathbf{b}_1=[\mathbf{I}+(1-C_1)\mathbf{M}+(1-C_1-C_2)\mathbf{M}^2+\cdots]\times[\mathbf{N}(\mathbf{Ba}_2)+\mathbf{S}_1(\mathbf{Ba}_1)]. \qquad (20)$$

It should be noted that $\mathbf{a}_1$ and $\mathbf{a}_2$ appear in (20) only in combination with $\mathbf{B}$. Equation (20) is part of the solution as it expresses the amplitudes of waves leaving the junction number (1) as a linear combination of the amplitudes of waves entering the combined network.

Since

$$\mathbf{S}_1\mathbf{Ba}_2=0 \qquad (21)$$

and

$$\mathbf{NBa}_1=0,$$

Eq. (20) can be written as

$$\mathbf{b}_1=[\mathbf{I}+(1-C_1)\mathbf{M}+(1-C_1-C_2)\mathbf{M}^2+\cdots]\times(\mathbf{N}+\mathbf{S}_1)\mathbf{B}(\mathbf{a}_1+\mathbf{a}_2). \qquad (22)$$

The amplitudes of waves leaving the junction number (2) remain to be calculated. This is easily done making use of (22). From (5) and (10)

$$\mathbf{b}_2=\mathbf{S}_2\mathbf{a}_2=\mathbf{S}_2(\mathbf{A}+\mathbf{B})\mathbf{a}_2=\mathbf{S}_2\mathbf{A}\mathbf{b}_1+\mathbf{S}_2\mathbf{Ba}_2. \qquad (23)$$

Substitute (22) in (23)

$$\mathbf{b}_2=\mathbf{S}_2\mathbf{A}[\mathbf{I}+(1-C_1)\mathbf{M}+(1-C_1-C_2)\mathbf{M}^2+\cdots]\times(\mathbf{N}+\mathbf{S}_1)\mathbf{B}(\mathbf{a}_1+\mathbf{a}_2)+\mathbf{S}_2\mathbf{Ba}_2, \qquad (24)$$

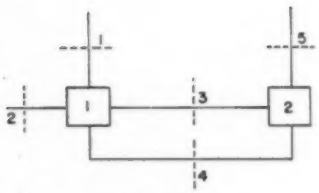


FIG. 3.

which may be written as

$$\mathbf{b}_2=\{\mathbf{S}_2+\mathbf{S}_2\mathbf{A}[\mathbf{I}+(1-C_1)\mathbf{M}+\cdots](\mathbf{N}+\mathbf{S}_1)\}\times\mathbf{B}(\mathbf{a}_1+a_2). \quad (25)$$

Equations (22) and (25) may be combined to form the scattering equation for the combined network.

$$\mathbf{B}(\mathbf{b}_1+\mathbf{b}_2)=\mathbf{S}\mathbf{B}(\mathbf{a}_1+\mathbf{a}_2) \quad (26)$$

where

$$\mathbf{S}=\mathbf{B}\{\mathbf{S}_2+(\mathbf{I}+\mathbf{S}_2\mathbf{A})[\mathbf{I}+(1-C_1)\mathbf{M}+(1-C_1-C_2)\mathbf{M}^2+\cdots](\mathbf{N}+\mathbf{S}_1)\}\mathbf{B}. \quad (27)$$

$\mathbf{S}$ as given by (27) is the scattering matrix of the combined network.

The network scattering matrix $\mathbf{S}$ has the following form:

$$\mathbf{S}=\left[\begin{array}{ccc|c} & 0 & & \\ 0 & 0 & 0 & r \\ & 0 & & \\ \hline \underbrace{\quad}_{p-r} & \underbrace{\quad}_{r} & \underbrace{\quad}_{q-r} & \end{array}\right]\begin{array}{l}\}p-r\\ \\ \}q-r\end{array}.$$

The r rows and columns which are zero may be removed to form a matrix of order $p+q-2r$. In a similar way the column vectors $\mathbf{B}(\mathbf{b}_1+\mathbf{b}_2)$ and $\mathbf{B}(\mathbf{a}_1+\mathbf{a}_2)$ have zeros in r rows beginning at the $p-r+1$ row. These zeros may be removed to form $p+q-2r$ rowed vectors. This process of removing zeros reduces the matrices to a dimension equal to the number of transmission lines feeding the combined network. The rows and columns removed are the ones associated with the interconnecting transmission lines.

For some purposes it is desirable to know the amplitudes of waves in the interconnecting transmission lines in terms of amplitudes of waves incident upon the network. These amplitudes are contained in $\mathbf{b}_1$ and $\mathbf{b}_2$ and may be expressed as

$$\mathbf{A}\mathbf{b}_1=\mathbf{A}[\mathbf{I}+(1-C_1)\mathbf{M}+(1-C_1-C_2)\mathbf{M}^2+\cdots]\times(\mathbf{N}+\mathbf{S}_1)\mathbf{B}(\mathbf{a}_1+\mathbf{a}_2) \quad (28)$$

and

$$\mathbf{A}\mathbf{b}_2=\mathbf{A}\{\mathbf{S}_2+\mathbf{S}_2\mathbf{A}[\mathbf{I}+(1-C_1)\mathbf{M}+\cdots]\times(\mathbf{N}+\mathbf{S}_1)\}\mathbf{B}(\mathbf{a}_1+\mathbf{a}_2). \quad (29)$$

The left-hand side of Eqs. (28) and (29) consists only of the amplitudes of waves in interconnecting transmission lines of the network.

To recapitulate, Eqs. (26)–(29) express the unknown complex wave amplitudes of waves in the transmission lines of a microwave network of the type illustrated in Fig. 2 as functions of the known amplitudes. In particular the left side of (26) consists of the amplitudes of waves leaving the network (along the lines 1, 2, 3, 7, and 8 of Fig. 2) expressed as a linear combination of the complex amplitudes, assumed known, of waves entering the network. In a similar way Eqs. (28) and (29) express the complex amplitudes of waves in the interconnecting transmission lines (4, 5, and 6, Fig. 2) as linear combinations of the known complex amplitudes of waves entering the network (along lines 1, 2, 3, 7, and 8 of Fig. 2).

EXAMPLE

As an example of the application of the formalism, consider the microwave network illustrated schematically in Fig. 3.

The "magic tee" has the property of dividing the power incident upon it in any one of its four feeding transmission lines equally between the two adjacent lines. For example, if a wave is sent in on line 2, it leaves the junction No. 1 *via* lines 1 and 4. Of course, the wave which leaves junction No. 1 *via* line 4 returns again after interacting with junction No. 2. This leads, in turn, to a new set of waves scattered by junction No. 1 and so on. By locating the terminals in (1), (2), (3), or (4) properly, the scattering matrix of the junction No. 1 may be written as

$$\mathbf{S}_1=\frac{1}{\sqrt{2}}\left[\begin{array}{cccc|c} 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & -1 & 0 \\ 0 & 1 & -1 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 \end{array}\right]. \quad (30)$$

As an example of the meaning of $\mathbf{S}_1$ note that the third column of $\mathbf{S}_1$ indicates that a wave of

amplitude 1 incident on the junction in transmission line 3 leaves the junction along lines 1 and 4 with amplitudes $1/\sqrt{2}$ and $-1/\sqrt{2}$, respectively. The phases of these waves are to be taken at their respective reference planes.

In a similar way the junction No. 2 is a matched symmetrical tee. Its scattering matrix is assumed to be

$$\mathbf{S}_2=\begin{bmatrix} 0 & 0 & 0 \\ 0 & \begin{matrix} 0 & 1/\sqrt{2} \\ 1/\sqrt{2} & \frac{1}{2} \end{matrix} & \begin{matrix} 1/\sqrt{2} \\ -\frac{1}{2} \end{matrix} \\ 0 & \begin{matrix} 1/\sqrt{2} & -\frac{1}{2} \end{matrix} & \frac{1}{2} \end{bmatrix}. \quad (31)$$

It is possible to write the scattering matrix of any matched symmetrical tee in this way merely by locating the terminals correctly. However, the terminals 3 (or 4) have already been fixed in position. Therefore, it is necessary to assume that either the line 3 or the line 4 has been adjusted to have the correct length.

It may be noted that both $\mathbf{S}_1$ and $\mathbf{S}_2$ are symmetric and unitary. These are general properties and are characteristic of all scattering matrices.

Turning now to the two auxiliary matrices $\mathbf{A}$ and $\mathbf{B}$, these have the following forms:

$$\mathbf{A}=\begin{bmatrix} 0 & 0 & 0 \\ 0 & \begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix} & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad (32)$$

$$\mathbf{B}=\begin{bmatrix} \begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix} & \begin{matrix} 0 \\ 0 \end{matrix} & \begin{matrix} 0 \\ 0 \end{matrix} \\ \begin{matrix} 0 & 0 \end{matrix} & 0 & 0 \\ \begin{matrix} 0 & 0 \end{matrix} & 0 & 1 \end{bmatrix}. \quad (33)$$

The two matrices $\mathbf{M}$ and $\mathbf{N}$ may now be computed (see Eq. (12)).

$$\mathbf{N}=\mathbf{S}_1\mathbf{A}\mathbf{S}_2$$

$$=\frac{1}{2\sqrt{2}}\begin{bmatrix} 0 & \begin{matrix} 0 & \sqrt{2} \\ \sqrt{2} & 1 \end{matrix} & \begin{matrix} \sqrt{2} \\ -1 \end{matrix} \\ 0 & \begin{matrix} -\sqrt{2} & -1 \\ 0 & -\sqrt{2} \end{matrix} & \begin{matrix} +1 \\ -\sqrt{2} \end{matrix} \\ 0 & 0 & 0 \end{bmatrix}, \quad (34)$$

$$\mathbf{M}=\mathbf{S}_1\mathbf{A}\mathbf{S}_2\mathbf{A}$$

$$=\frac{1}{2\sqrt{2}}\begin{bmatrix} 0 & \begin{matrix} 0 & \sqrt{2} \\ \sqrt{2} & 1 \end{matrix} & 0 \\ 0 & \begin{matrix} -\sqrt{2} & -1 \\ 0 & -\sqrt{2} \end{matrix} & 0 \\ 0 & 0 & 0 \end{bmatrix}. \quad (35)$$

Equation (16) may be obtained by using the method introduced in the appendix [see Eq. (45)]

$$0=\sum_n C_n\lambda^n=\lambda\det(\mathbf{R}-\mathbf{I}\lambda)=\lambda\begin{vmatrix} -(\lambda+\frac{1}{2}) & \frac{-1}{2\sqrt{2}} \\ 0 & -(\lambda+\frac{1}{2}) \end{vmatrix}$$

$$=\lambda(\lambda+\tfrac{1}{2})^2=\lambda^3+\lambda^2+\tfrac{1}{4}\lambda. \quad (36)$$

The coefficients C_n must be normalized to satisfy Eq. (18). They then have the values

$$C_1=\frac{1}{9}, \quad C_2=\frac{4}{9}, \quad C_3=\frac{4}{9}. \quad (37)$$

The scattering matrix of the combined network may be computed by substituting in Eq. (27)

$$\mathbf{S}=\mathbf{B}\left\{\mathbf{S}_2+(\mathbf{I}+\mathbf{S}_2\mathbf{A})\left[\mathbf{I}+\tfrac{8}{9}\mathbf{M}+\tfrac{5}{9}\mathbf{M}^2\right](\mathbf{N}+\mathbf{S}_1)\right\}\mathbf{B}$$

$$=\tfrac{1}{3}\begin{bmatrix} 0 & 2\sqrt{2} & \begin{matrix} 0 \\ \end{matrix} & 1 \\ \begin{matrix} 2\sqrt{2} & \frac{1}{3} \end{matrix} & & \begin{matrix} -\frac{2\sqrt{2}}{3} \end{matrix} \\ \begin{matrix} 0 \end{matrix} & \begin{matrix} 0 \end{matrix} & \begin{matrix} 0 \end{matrix} \\ \begin{matrix} 1-\frac{2\sqrt{2}}{3} \end{matrix} & \begin{matrix} 0 \end{matrix} & \begin{matrix} \frac{8}{3} \end{matrix} \end{bmatrix}. \quad (38)$$

If the two rows and columns which vanish are removed, the matrix becomes

$$\mathbf{S}=\tfrac{1}{3}\begin{bmatrix} 0 & 2\sqrt{2} & 1 \\ 2\sqrt{2} & \frac{1}{3} & -\frac{2\sqrt{2}}{3} \\ 1 & -\frac{2\sqrt{2}}{3} & \frac{8}{3} \end{bmatrix}. \quad (39)$$

It should be noted that this matrix is symmetric and unitary as it should be.

The matrices which can be used to calculate the amplitudes of waves in the connecting lines are obtained from Eqs. (28) and (29).

$$\mathbf{S}_c=\mathbf{A}\left[\mathbf{I}+\frac{8}{9}\mathbf{M}+\frac{4}{9}\mathbf{M}^2\right](\mathbf{N}+\mathbf{S}_1)\mathbf{B}$$

$$=\frac{1}{3}\begin{bmatrix} 0 & 0 & 0 \\ \sqrt{2} & -\frac{1}{3} & \frac{2\sqrt{2}}{3} \\ 0 & \sqrt{2} & -1 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\mathbf{S}_c'=\mathbf{A}\left\{\mathbf{S}_2+\mathbf{S}_2\mathbf{A}\left[\mathbf{I}+\frac{8}{9}\mathbf{M}+\frac{4}{9}\mathbf{M}^2\right](\mathbf{N}+\mathbf{S}_1)\right\}\mathbf{B}$$

$$=-\frac{1}{9}\begin{bmatrix} 0 & 0 & 0 \\ 0 & 3 & 3\sqrt{2} \\ 3 & \sqrt{2} & -4 \\ 0 & 0 & 0 \end{bmatrix}. \quad (41)$$

To illustrate the meaning of the matrices (38), (40), and (41), assume that a wave of unit amplitude is incident upon the combined network at terminals No. 1. The scattering matrix **S**, Eq. (39), states that the amplitudes of the waves leaving the network *via* lines (2) and (5) are $\frac{2}{3}\sqrt{2}$ and $\frac{1}{3}$, respectively. The matrix (40) states that there is a wave of amplitude $\sqrt{2}/3$ running in line (3) from junction No. 1 to No. 2. There is no wave moving in this direction along line (4). The matrix (41) states that there is no wave moving from junction No. 2 to No. 1 along (3), but that there is a wave of amplitude $\frac{1}{3}$ moving from right to left in line (4).

APPENDIX

From Eqs. (3), (6), and (12) it is apparent that **M** has the form

$$\mathbf{M}=\begin{bmatrix} 0 & - & 0 \\ 0 & \mathbf{R} & 0 \\ 0 & 0 & 0 \\ \underbrace{}_{p-r} & \underbrace{}_{r} & \underbrace{}_{q-r} \end{bmatrix}\begin{matrix} \}p-r \\ \}r \\ \}q-r \end{matrix}. \quad (42)$$

The characteristic equation of **M** is defined as

$$\det(\mathbf{M}-\lambda\mathbf{I})=0. \quad (43)$$

Denote the middle square of **M** as a sub-matrix **R**. It is evident that

$$\det(\mathbf{M}-\lambda\mathbf{I})=\lambda^{q+p-2r}\det(\mathbf{R}-\mathbf{I}\lambda), \quad (44)$$

where the determinant on the right side of (32) is of order r. In order to reduce the characteristic polynomial to its lowest degree, Eq. (16) is defined by the relation

$$0=\sum_n C_n\lambda^n=\lambda\det(\mathbf{R}-\mathbf{I}\lambda). \quad (45)$$

Thus the characteristic polynomial is of degree $r+1$ and the index n runs from 1 to $r+1$.

It will be remembered that the above described computational method depended for its validity on the non-vanishing of the sum of the coefficients of (33). A necessary and sufficient condition for the vanishing of this sum is the occurrence of

$$\lambda=1$$

as an eigenvalue of the matrix **M**. The values of the eigenvalues of **M** depend upon the scattering matrices $\mathbf{S}_1$ and $\mathbf{S}_2$. However, the locations of the reference planes or terminals in the interconnecting transmission lines are arbitrary. By changing their locations an accidental occurence of the root $\lambda=1$ can be eliminated.

The Optics of Three-Electrode Electron Guns*

S. G. Ellis
The McLennan Laboratory, Department of Physics, University of Toronto, Toronto, Canada
(Received April 17, 1947)

The paper is concerned primarily with the electron guns employed in electron microscopy. The position of the image of the cathode, the magnification, and the divergence of the beam leaving the anode are calculated for a three-electrode electron gun with plane electrodes and circular apertures both for zero and battery bias. The simplifying assumption is made that the electrostatic lenses are thin. The results are compared with the experiments of Johannson. The extension of the results to the gun used in the transmission electron microscope is considered with the aid of electrolytic trough measurements. The case of a gun biased by a cathode resistor is also discussed.

INTRODUCTION

THE illuminating system of the transmission electron microscope was first treated analytically by von Borries and Ruska.[1] They considered the electron gun as a source of electrons of fixed cross section and position and considered the dependence of the semi-angle of the illuminating beam at the specimen on the strength of the condenser lens. Marton and Hutter[2] extended this treatment to include the converging effect on the illuminating beam of that part of the objective field which lies between the specimen and the condenser lens. Finally, Hillier and Baker[3] have considered, semi-empirically, the influence of the filament height on the angular aperture of the illumination by treating the electron gun as a converging lens containing a physical aperture which may itself limit the aperture of the illumination.

Hitherto the electron gun has received either very general physical treatment,[4] or rather special semi-empirical treatment[5] in the literature. Since these treatments do not greatly clarify the behavior of the particular types of electron gun employed in electron microscopy, an attempt has been made in this paper to develop an analytic theory of a very simplified electron gun. The main simplifications are the treatment of the electrodes as plane, and the variations of axial potential as linear. The latter simplification is equivalent to treating the lenses as thin lenses. Later, this theory is used, in conjunction with electrolytic trough measurements, to describe the behavior of the guns employed in practice.

This analytic treatment is aided by a theorem attributed to Schlesinger[6] which may be called the "Transit Time Theorem." Consider an electron leaving an object (cathode or image) at A, Fig. 1, passing through a lens at B and being focused to C. An axis Oz is taken, coaxial with the lens, positive in the direction of electron motion and with origin at the lens. Take the zero of time as the epoch at which the electron passes through the lens OB. T_0 is the epoch at which the electron leaves A and T_I the epoch at which it arrives at C. It is supposed that the lens is thin so that radial forces act on the electron only while it is passing the OB plane. If then a radial impulse I, positive if centrifugal, acts on the electron when passing through the lens at a distance R from the center, then Schlesinger shows that the condition for image formation is:

$$I \propto R \tag{1}$$

and that under these conditions

$$(1/T_I)-(1/T_0)=-(I/mR) \tag{2}$$

$$=1/T_F \text{ say.} \tag{3}$$

In Eq. (2) m is the mass of the electron. Equation (2) immediately leads to Gaussian dioptrics in the magnetic lens and the bipotential lens.

* Based on Section A of a thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy in the University of Toronto. Part of this work was supported by the grant of a Fellowship by the National Research Council, Ottawa, Canada,

[1] B. von Borries and E. Ruska, Zeits. f. tech. Physik. **20**, 225 (1939).
[2] L. Marton and R. C. E. Hutter, Phys. Rev. **65**, 161 (1944).
[3] J. Hillier and R. F. Baker, J. App. Phys. **16**, 469 (1945).
[4] L. M. Field, Rev. Mod. Phys. **18**, 353 (1946).
[5] G. A. Morton, Rev. Mod. Phys. **18**, 362 (1946).
[6] K. Schlesinger, Proc. I.R.E. **32**, 483 (1944).

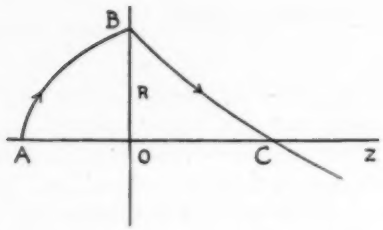


FIG. 1. General electron trajectory through a thin lens.

where the axial velocity of the electron is constant outside the lens and the trajectories are linear there. In the immersion objective the dioptrics will not be Gaussian, both because the axial velocity is continually changing and also because the strength of the lens (proportional to I/R) will depend on the object distance OA. The magnification, M, for the thin lens is, quite generally, given by

$$M = T_I/T_0. \tag{4}$$

SIMPLIFYING ASSUMPTIONS AND DEFINITIONS

The filament or cathode is considered as a plane circular emitting area of radius r_s (see Fig. 2). The shield, or grid, is considered as a thin sheet with a central aperture of radius r_k'', distant d_2 from a plane anode. An axis $O_s z$ with origin O_s at the center of the shield aperture and positive in the direction of electron motion is chosen to coincide with the axis of symmetry of the electron gun. The assumed variation of axial potential is shown in Fig. 3.

The following factors have now to be taken into consideration. The electrons leave the filament with a velocity distribution both in magnitude and direction governed by the temperature of the filament. The electrostatic field in the vicinity of the shield exerts a convergent effect upon the electrons. The electron speeds are nowhere constant within the electron gun, and it is therefore convenient to use the transit time method of Schlesinger taking zero time as the instant at which the electron passes through the shield.

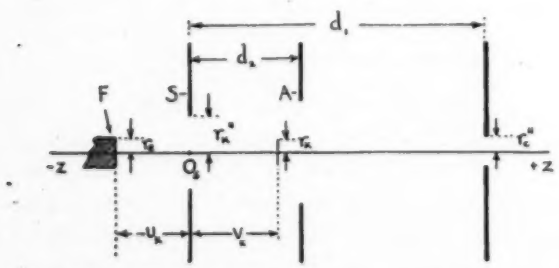


FIG. 2. Symbols used in text. F represents the filament, S the shield, and A the anode.

To commence, consider an electron which leaves the filament at the axis with zero axial velocity and with transverse thermal velocity v_T cm per sec. (or V_T in electron volts). It describes a parabolic path (Fig. 4) until it reaches the shield aperture at a distance R from the axis. At this point it receives a centripetal impulse I_k and returns to the axis at a distance v_k from the shield, following another parabola until it reaches the anode aperture where it receives a centrifugal impulse I_A traveling thereafter in a straight line towards the condenser lens of the microscope.

NOTATION

Where possible the notation employed by Hillier and Baker[3] has been used in this section.

r_s = radius of emitting area, assumed plane
r_k'' = radius of hole in shield
r_c'' = radius of condenser aperture
r_0' = radius of area of specimen plane illuminated by electrons
V_k = potential at origin O_s
V_k' = potential of shield
V = potential of anode
V_T = transverse (thermal) velocity of electron leaving filament in electron volts
E_1 = electrostatic field strength between filament and shield = V_k/u_k
E_2 = electrostatic field strength between shield and anode = $-V/d_2$ (if $V_k \ll V$)
v_T = transverse (thermal) velocity of electron leaving filament in cm per sec.
T_0 = epoch at which electron leaves filament (negative with sign convention used here)
T_I = epoch at which electron arrives at real image formed by cathode lens
T_{SA} = transit time of electron from shield to anode
T_0' = epoch at which electron leaves real image formed by cathode lens
T_I' = epoch at which electron 'passes' virtual image formed by anode aperture
u_k = distance from shield to filament (negative on above sign convention) commonly called the filament height
v_k = distance from shield to image formed by cathode lens
d_1 = distance from shield to condenser aperture
d_2 = distance from shield to anode
R = radial displacement of electron at shield
$\dot{R}$ = radial velocity of electron between shield and anode

R_A = radial displacement of electron at anode
$I_k = I_k(R)$ = centripetal impulse acting on electron on passing shield
$I_A = I_A(R_A)$ = centrifugal impulse acting on electron on passing anode
α_c max. = maximum angular aperture of illumination at specimen
β = angular aperture of beam at cathode side of anode
β' = angular aperture of beam leaving the electron gun
γ = angle between $O_s z$ and center of beam from off-axial point of cathode, see Fig. 8
M_k = magnification produced by cathode lens
M_A = magnification produced by anode lens
M_G = over-all magnification produced by electron gun
M_c = magnification produced by condenser lens
e = charge of electron
m = mass of electron.

THEORY

From the transit time theory of Schlesinger

$$(1/T_I)-(1/T_0)=-(I_k/mR)=1/T_F. \quad (5)$$

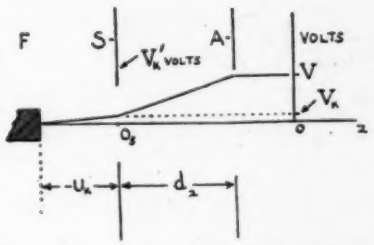


FIG. 3. Potential distribution assumed.

Further

$$-u_k=\frac{1}{2}\left(\frac{eV_k}{mu_k}\right)T_0^2. \quad (6)$$

At the plane of the shield the electron will have gained an axial component of velocity $(-2eV_k/m)^{\frac{1}{2}}$. Hence

$$v_k=\left(-\frac{2eV_k}{m}\right)^{\frac{1}{2}}T_I-\frac{1}{2}\cdot\frac{eV}{md_2}\cdot T_I^2 \quad (7)$$

and

$$T_I=\frac{2(-Ve/2md_2)^{\frac{1}{2}}\cdot[\lambda(1-\epsilon)]^{\frac{1}{2}}\pm\{4(-Ve/2md_2)\cdot[\lambda(1-\epsilon)]+4(-Ve/2md_2)v_k\}^{\frac{1}{2}}}{-2(-Ve/2md_2)}, \quad (8)$$

where

$$\lambda=u_k r_k''/(\pi u_k-r_k'') \quad (9)$$

and

$$\epsilon=-\pi(d_2/r_k'')\cdot(V_k'/V). \quad (10)$$

Then with an obvious choice of sign

$$T_I=\{[\lambda(1-\epsilon)+v_k]^{\frac{1}{2}}-[\lambda(1-\epsilon)]^{\frac{1}{2}}\}/(-Ve/2md_2)^{\frac{1}{2}}. \quad (11)$$

To obtain the value of I_k/mR notice that if the lens may be treated as a thin lens[7]

$$\frac{I_k}{mR}=-\left(-\frac{2V_k e}{m}\right)^{\frac{1}{2}}\cdot\left[\frac{E_2-E_1}{4V_k}\right], \quad (12)$$

but[8]

$$V_k=V\lambda(1-\epsilon)/d_2 \quad (13)$$

and, therefore,

$$\frac{I_k}{mR}=\left[-\frac{Ve}{2md_2}\right]^{\frac{1}{2}}\cdot\left[\frac{1}{2[\lambda(1-\epsilon)]^{\frac{1}{2}}}+\frac{[\lambda(1-\epsilon)]^{\frac{1}{2}}}{2u_k}\right]. \quad (14)$$

From Eqs. (6), (11), and (14)

$$\frac{1}{T_0}=\left(-\frac{Ve}{2md_2}\right)^{\frac{1}{2}}\cdot\frac{1}{u_k}\cdot[\lambda(1-\epsilon)]^{\frac{1}{2}} \quad (15)$$

$$\frac{1}{T_I}=\left(-\frac{Ve}{2md_2}\right)^{\frac{1}{2}}\cdot\frac{1}{[\lambda(1-\epsilon)+v_k]^{\frac{1}{2}}-[\lambda(1-\epsilon)]^{\frac{1}{2}}} \quad (16)$$

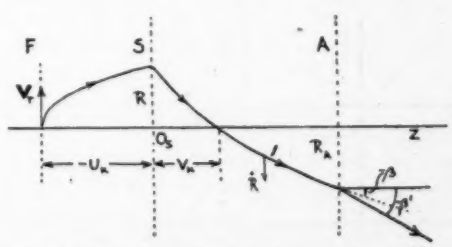


FIG. 4. An electron trajectory for the three-electrode electron gun.

[7] Zworykin, Morton, Ramberg, Hillier, and Vance, *Electron Optics and the Electron Microscope* (John Wiley and Sons, Inc., New York, 1945), p. 444.
[8] Zworykin, Morton, Ramberg, Hillier, and Vance, *Electron Optics on the Electron Microscope* (John Wiley and Sons, Inc., New York, 1945), pp. 383–385, particularly equation 11.95.

$$\frac{1}{T_F}=\left(-\frac{Ve}{2md_2}\right)^{\frac{1}{2}}\cdot\left\{[\lambda(1-\epsilon)]^{\frac{1}{2}}\left[\frac{1}{2\lambda(1-\epsilon)}+\frac{1}{2u_k}\right]\right\}. \quad (17)$$

Substituting from Eqs. (15) to (17) into Eq. (5) and solving for v_k

$$v_k=4[\lambda(1-\epsilon)]\left\{\frac{1+(1+3[\lambda(1-\epsilon)]/u_k)}{(1+3[\lambda(1-\epsilon)]/u_k)^2}\right\}. \quad (18)$$

Values of v_k are plotted in Fig. 5 for typical values of the other variables. The non-Gaussian dioptrics of the system is apparent from Eq. (18).

DIVERGENCE OF THE BEAM

One of the factors controlling the intensity of illumination in the electron microscope will be the divergence of the electron beam leaving the electron gun, for this will determine the percentage of electrons leaving the electron gun which pass through the condenser diaphragm.

If β is the angle to the axis O_0z at which the peripheral beam approaches the anode then, from Fig. 4

$$\tan\beta=\frac{\dot{R}}{(-2eV/m)^{\frac{1}{2}}}=\frac{I_k/m+v_T}{(-2eV/m)^{\frac{1}{2}}}, \quad (19)$$

but

$$I_k/m=-R/T_F \quad (20)$$

and

$$R=-v_T T_0; \quad (21)$$

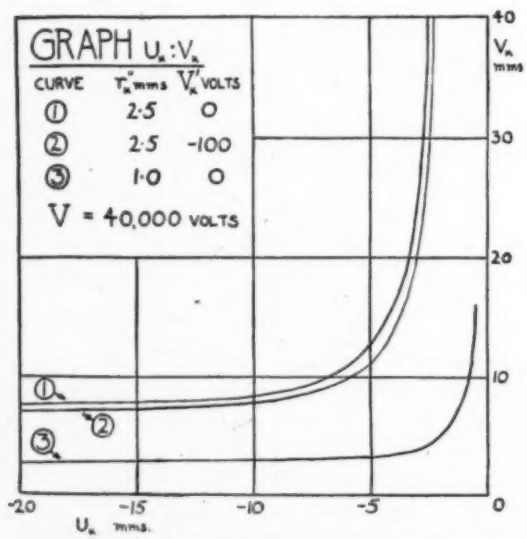


FIG. 5.

therefore

$$\tan\beta=\left(\frac{V_T}{V}\right)^{\frac{1}{2}}\left[1+\frac{T_0}{T_F}\right]. \quad (22)$$

Substituting in Eq. (22) from Eqs. (15) and (17)

$$\tan\beta=\frac{3}{2}\left(\frac{V_T}{V}\right)^{\frac{1}{2}}\left[1-\frac{1}{3(1-\epsilon)}+\frac{\pi u_k}{3r_k''(1-\epsilon)}\right]. \quad (23)$$

The further divergence produced in the electron beam by the change in field strength at the anode aperture may now be determined. Then

$$\tan\beta'=\frac{\dot{R}+I_A/m}{(-2eV/m)^{\frac{1}{2}}} \quad (24)$$

$$=\tan\beta+I_A/m(-2eV/m)^{\frac{1}{2}}. \quad (25)$$

But,[7]

$$\frac{I_A}{mR_A}=\left(-\frac{2eV}{m}\right)^{\frac{1}{2}}\cdot\frac{1}{4d_2} \quad (26)$$

$$\frac{I_A}{m(-2eV/m)^{\frac{1}{2}}}=\frac{R_A}{4d_2} \quad (27)$$

and

$$\tan\beta'=\tan\beta+R_A/4d_2. \quad (28)$$

Further

$$R+R_A=\dot{R}T_{SA} \quad (29)$$

and

$$T_{SA}=T_I \text{ (for } v_k=d_2), \quad (30)$$

that is

$$T_{SA}=(-2d_2m/Ve)^{\frac{1}{2}}\{[\lambda(1-\epsilon)+d_2]^{\frac{1}{2}}-[\lambda(1-\epsilon)]^{\frac{1}{2}}\} \quad (31)$$

and approximating

$$T_{SA}=(-2d_2m/Ve)^{\frac{1}{2}}\cdot d_2^{\frac{1}{2}} \quad (32)$$

giving

$$R_A=2d_2\tan\beta+2\left(\frac{V_T}{V}\right)^{\frac{1}{2}}\cdot\frac{d_2^{\frac{1}{2}}u_k}{[\lambda(1-\epsilon)]^{\frac{1}{2}}} \quad (33)$$

and

$$\tan\beta'=\frac{3}{2}\tan\beta+\frac{1}{2}\left(\frac{V_T}{V}\right)^{\frac{1}{2}}\frac{u_k}{[\lambda d_2(1-\epsilon)]^{\frac{1}{2}}}. \quad (34)$$

The negative values of β' obtained from Eq. (34) indicate that the electrons from the filament cross the axis of the gun. Values of $\tan\beta'$ are plotted in Fig. 6. It has been assumed here that R is less than the radius of the $-V_T$ equipotential surfaces at the cathode shield, which is true for the range of u_k plotted in the graphs.

MAGNIFICATION

It would appear from Fig. 5 that for the range of filament heights commonly employed in the electron microscope, the cathode lens forms a real image of the filament between the shield and the anode (i.e., $v_k<d_2$). First, therefore, the magnification M_k produced by the cathode lens may be calculated.

By the transit time method of Schlesinger

$$M_k=T_I/T_0. \tag{35}$$

Negative values obtained for M_k indicate inversion of the image with respect to the object. Using Eqs. (15) and (16)

$$M_k=[\lambda(1-\epsilon)]^{\frac{1}{2}}\{[\lambda(1-\epsilon)+v_k]^{\frac{1}{2}}-[\lambda(1-\epsilon)]^{\frac{1}{2}}\}/u_k. \tag{36}$$

The anode lens forms a virtual image of the real image mentioned above. For $v_k<d_2$ this virtual image lies between v_k and d_2 and is demagnified. If M_A is the magnification produced by the anode lens alone then

$$M_A=T_I'/T_0' \tag{37}$$

but

$$(1/T_I')-(1/T_0')=-(I_A/mR_A) \tag{38}$$

$$=1/T_F' \text{ say} \tag{39}$$

so that

$$M_A=T_F'/(T_0'+T_F') \tag{40}$$

writing

$$T_0'=-(T_{SA}-T_I) \tag{41}$$

and using Eqs. (31) and (16)

$$M_A=\frac{2d_2^{\frac{1}{2}}\{[\lambda(1-\epsilon)+v_k]^{\frac{1}{2}}+[\lambda(1-\epsilon)+d_2]^{\frac{1}{2}}\}}{(d_2-v_k)+2d_2^{\frac{1}{2}}\{[\lambda(1-\epsilon)+v_k]^{\frac{1}{2}}+[\lambda(1-\epsilon)+d_2]^{\frac{1}{2}}\}}. \tag{42}$$

Values of M_k and M_A are plotted against u_k in Fig. 7.

AREA OF SPECIMEN ILLUMINATED IN THE ELECTRON MICROSCOPE

On the basis of the foregoing theory one might expect the radius of the illuminated part of the specimen in an electron microscope r_0' to be given by

$$r_0'=r_eM_kM_AM_c, \tag{43}$$

where M_c is the magnification produced by the condenser lens. Writing

$$M_kM_A=M_G \tag{44}$$

the magnification produced by the electron gun, and taking $M_c=1$ then

$$r_0'=r_eM_G. \tag{45}$$

Equation (45) applies only to the case in which the virtual image of the cathode due to the anode aperture is focused on to the specimen by the condenser lens.

For values of u_k that make $\tan\beta'$ small, another factor may have a decisive effect on the value of r_0'. The condenser aperture may act as a field limiting aperture for the electron gun even though the image produced by the gun is not in the plane of the condenser aperture.

It will be seen from Fig. 8 that electrons leaving a point in the cathode distant r_e' from the axis leave the electron gun in a cone of semi-angle β', the axis of the cone making an angle γ with the axis of the system. If the radius of the condenser lens diaphragm is r_c'', and it is situated a distance d_1 from O_s, the condition that electrons from P shall pass the condenser lens is then, approximately

$$(\gamma-\beta')(d_1-v_k)\leqslant r_c''-|M_G|r_e' \tag{46}$$

with

$$r_e'\leqslant r_c. \tag{47}$$

It is desirable, therefore, to investigate the dependence of γ upon r_e'. Considering the ray Pcc' and employing the methods of the preceding sections, it can be shown that

$$\gamma=\frac{r_e'}{4}\left\{\left[\frac{\lambda(1-\epsilon)}{d_2}\right]^{\frac{1}{2}}\left[\frac{1}{\lambda(1-\epsilon)}+\frac{1}{u_k}\right]+\frac{1}{d_2}\right\}. \tag{48}$$

Neglecting the term $|M_G|r_e'$ in Eq. (46), and taking

$$r_c''/(d_1-v_k)=\alpha_c \text{ max}, \tag{49}$$

we can calculate the radius r_e'' of the area of the cathode from which electrons can pass the condenser aperture. Equation (45) presumes that the electron gun and the specimen are equidistant from the condenser lens. Then from Eqs. (46),

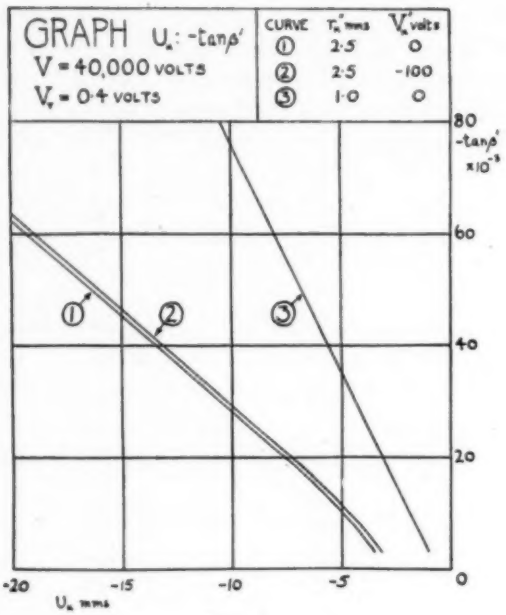


FIG. 6.

(48), and (49)

$$r_e'' = 4(\alpha_c \max + |\beta'|) \Big/ \left\{ \left[\frac{\lambda(1-\epsilon)}{d_2} \right]^{\frac{1}{2}} \times \left[\frac{1}{\lambda(1-\epsilon)} + \frac{1}{u_k} \right] + \frac{1}{d_2} \right\} \quad (50)$$

$$r_0' = r_e'' M_G. \quad (51)$$

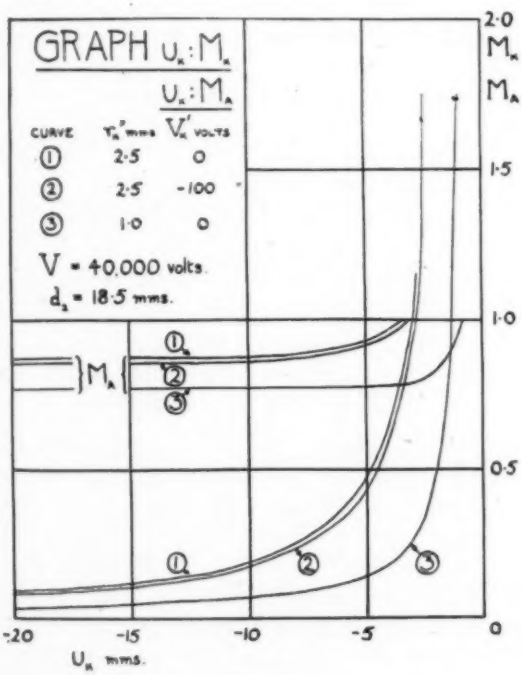


FIG. 7.

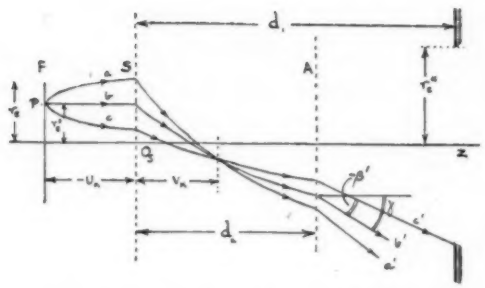


FIG. 8. Trajectories in beam from off-axial point in the filament.

In practice r_0' will be given by the smaller of the values in Eqs. (45) or (51). In Fig. 9 r_0' is plotted against u_k the full line showing the variation from Eq. (45), and the broken line the variation from Eq. (51) for a sufficiently large cathode that relation 47 is maintained at all u_k. (The corresponding curves for conditions 1 and 2 lie off the graph with the scale chosen).

It has sometimes been stated that the smallness of the effective source provided by the electron gun is due to the presence of a cross-over of smaller radius than the image of the cathode. From Fig. 10 it can be seen that such a cross-over exists only if $\beta' < \gamma$. This is approximately the condition for Eq. (51) to limit the area of illumination of the specimen.

On the basis of the foregoing calculation one might conclude that the source in an electron microscope operated with a zero-biased, or battery-biased gun and with the filament at a normal height is a reduced image of the filament and not a cross-over. This will be discussed later.

One other point should be mentioned. Only the

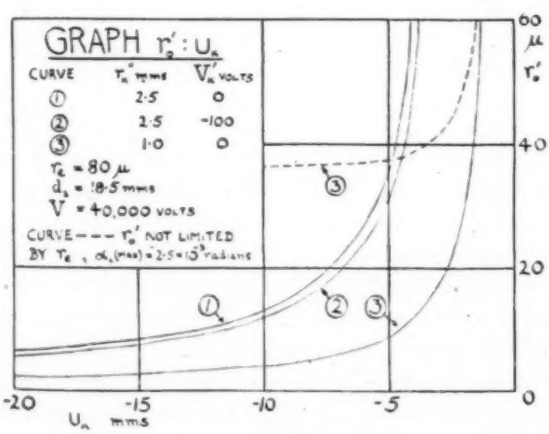


FIG. 9.

trajectories of electrons which leave the filament with zero axial velocity have been considered above. The extension of the theory to cover those electrons which have axial velocities V_T (in electron volts) at the filament would not be difficult but is not attempted here, since the above treatment suffices for the purpose in view. We may note in passing that for the electron gun as commonly employed in the electron microscope $V_T \ll V$, so that the longitudinal chromatic aberration as calculated on the above theory would not be very large.

THE IDEAL ELECTRON GUN

The heating effect of the beam acting on the specimen is proportional to $\sigma r_0'^2$, where σ is the current density at the specimen. The cross-sectional area of film over which this heat can be dissipated is proportional to r_0'. There is an obvious advantage in reducing r_0'. To a first approximation if r_0' is halved, σ may be made four times greater for the same rise in temperature of the specimen. Ideally one would prefer to illuminate only that part of the specimen which is examined on the final screen. If other parts of the specimen are illuminated scattered electrons from these regions may reach the final screen and reduce the contrast.

The illuminating beam, if it does not strike any part of the supporting mesh, will build up a static charge on the film and specimen. Both electron microscopy and electron diffraction require an electron gun giving a small, intense, single source of electrons.

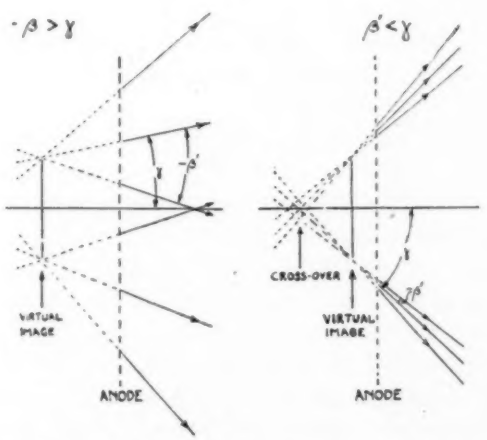


FIG. 10. Conditions for the production of a cross-over.

The angular aperture, α_c, of the illumination governs the depth of focus and intensity at constant beam current. When an electron microscope is used for electron diffraction with the specimen below the projector lens the radius of the objective exit pupil which acts as the effective source for the diffraction camera is governed by the aperture of the illumination. The aperture, α_c, will influence the resolving power of the microscope, particularly if the objective field lacks symmetry, for in this case when α_0 is large the rays leaving the object will move into those parts of the objective field further from the axis with a consequent increase in image assymetries. It has been argued, though convincing experimental proof is lacking, that even with a symmetrical objective field there will be an optimum value for α_0 which will be attained for a certain value of α_c.

Other factors to be considered in the choice of an electron gun are the ease with which it can be assembled and cleaned. Both the electron gun and condenser lens should form a system which can be readily adjusted to provide an illuminating beam coaxial with the objective lens.

Ideally then one would prefer to have r_0', σ, and α_c independently variable. In practice, since simplicity of construction is very desirable, this ideal is sacrificed and the chief aim is to provide a small source of adequate intensity.

PRACTICAL ELECTRON GUNS

The electron guns employed in practice may be divided into two classes according to the conditions under which they are operated. These classes are, the fixed-bias gun operated at emission saturation, and the cathode-biased gun[9]

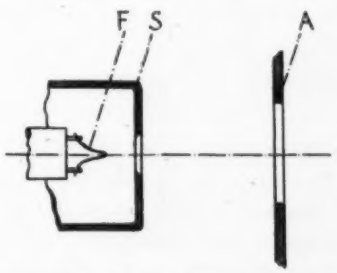


FIG. 11. A three-electrode electron gun as employed in electron microscopy.

[9] J. Hillier and R. F. Baker, J. App. Phys. **17**, 12 (1946).

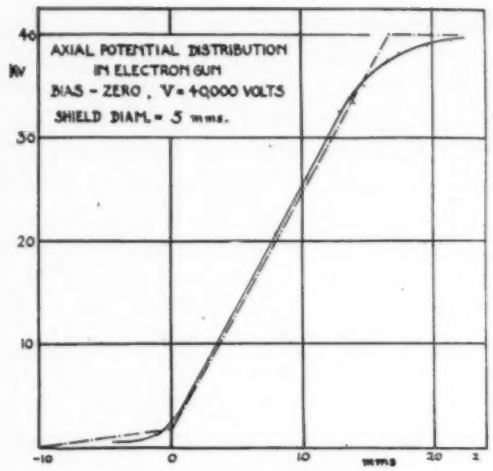


FIG. 12.

operated with space charge limited current.[8] This is not intended to imply that space charge is without effect in the former class of electron gun.

The construction of a fixed-bias gun is shown in Fig. 11. *F* is a hairpin filament of tungsten heated by a d.c. or r-f current. *S* is the shield or grid. It may be maintained at the same potential as the filament or at a negative potential of up to 200 volts with respect to the filament by a battery. *A* is the plate or anode which will be at a voltage of from 30 to 200 kv with respect to the filament in electron microscope practice. Matters of practical interest in the operation of such a gun have been described by Hillier and Baker.[3] Our immediate concern is to see wherein the optical behavior of this type of gun departs from the simple theory outlined above.

The major limitations in the above theory are imposed by the assumption of plane electrodes and linear axial potential variation. With this in mind experiments were undertaken with an electrolytic trough to determine the potential distributions occurring in practice.

Two points may be mentioned in connection with the use of the electrolytic trough. First, since the trough is a relatively large and open structure, it is difficult to shield it from the effect of stray fields which may become troublesome if measurements of high sensitivity are attempted. It is therefore desirable to use for the generator a supply of variable audiofrequency and for the detector a pair of head phones so that the balance can be made to the generator note and stray noises ignored. This is particularly important if an amplifier is used between the probe and the detector, and might well lead to errors if a vacuum-tube voltmeter were employed.

The second point concerns the plotting of the equipotential lines close to the filament. The field distribution near the filament will have a considerable influence on the behavior of the gun since the electrons are moving at slow speeds in this region. It is difficult to determine the equipotentials in this region, however, since the arms of the potential divider have to be made very unequal and the sensitivity of the circuit (which may be considered as a Wheatstone net) becomes very small. Under these conditions the sensitivity was improved by turning a copper plate to the shape of a fairly flat equipotential surface just outside the shield, placing this plate in the tank in the appropriate position, and transfering the lead from the anode to this shaped

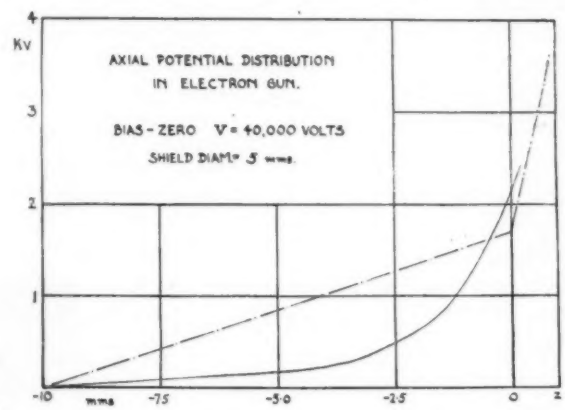


FIG. 13.

plate. The low equipotentials could then be studied with less unequal values of the arms of the potential divider.

Figure 12 shows the axial potential distribution for an electron gun of the form shown in Fig. 11. The full line is from the experimental results; the broken line shows the distribution assumed in the elementary theory above. The origin has been taken at the shield center. In Fig. 13 the distribution of axial potential between the shield and the filament is shown in more detail for $u_k = -10$ mm. Measurements were made with the plane of the filament both in and perpendicular to the surface of the electrolyte. No very significant

differences in the axial potential variation were detected in the two cases. It is probable that the magnetic effect of the heating current in the filament would produce effects at least as great as those produced by assymetries in the electric field. For values of z between -5 mm and -10 mm the measured potentials were of doubtful accuracy, and the curve has been extrapolated linearly to $z=-10$ mm. When u_k was reduced to -5.0 mm (Fig. 14) it could be observed that the curve for the axial potential distribution was concave downwards in the vicinity of the filament. In other words, the field strength is relatively high at the filament because of the small radii of curvature of its surface. Under these conditions the optical description of the electron gun is as follows. Leaving the filament the electrons first pass through a thick diverging lens, then a thick converging lens which terminates shortly beyond the shield; at the anode the electron passes through a thick diverging lens. The net effect of these lenses, as compared with the thin lenses treated above will be to further reduce M_G, the gun magnification, and to increase the magnitude of β', the divergence of the beam from any point of the filament. Under normal conditions of operation, therefore, a cross-over will not be produced in this electron gun.

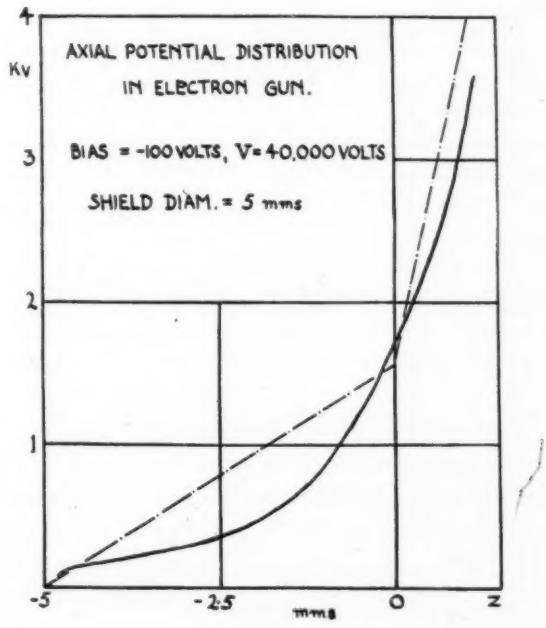


FIG. 14.

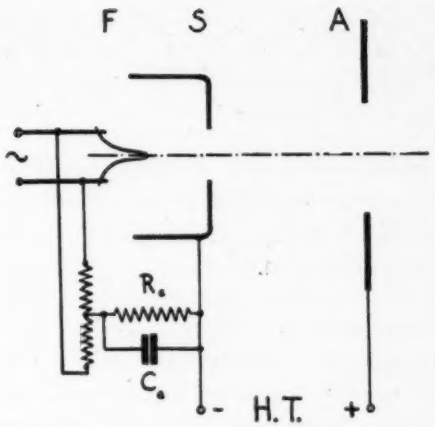


FIG. 15. The cathode-biased electron gun.

Observations with electron microscopes do not discredit this view. In the electron microscope put into operation in 1944 at the University of Toronto[10] the condenser diaphragm can be displaced laterally while the microscope is in operation. If this is done, the region of illumination of the specimen as observed at the intermediate image is seen to move. While this at first sight might seem to indicate the state of affairs shown in Fig. 8, and the existence of a cross-over, it is considered that the effect is due largely to the spherical aberration of the condenser lens.

The major disadvantages of this type of electron gun are the multiple nature of the source, the large source size which must be employed if high intensity is required, and the consequent large beam current. The gun is also inefficient because the large values of β' employed mean that most of the electrons hit the condenser diaphragm. The analysis given above indicates a fundamental difficulty in the design of this type of electron gun. The condition for high intensity is β' small, which from Eq. 22 indicates

$$\left(\frac{V_T}{V}\right)^{\frac{1}{2}} \cdot \frac{T_0+T_F}{T_F} \quad \text{small.}$$

The condition for a small area of illumination of the specimen is that M_G be small which requires

$$\frac{T_0+T_F}{T_F} \quad \text{to be large.}$$

[10] E. F. Burton and W. H. Kohl, *The Electron Microscope* (Reinhold Publishing Corporation, New York, 1946), second edition, Chap. 18.

Unless V_T/V can be reduced or the area of illumination of the specimen independently controlled the final design and operating condition must be a compromise.

CATHODE-BIASED ELECTRON GUNS

The essential parts of a cathode-biased electron gun are shown in Fig. 15. For small filament and hence small beam currents its optical behavior is similar to that of the fixed-bias gun. As the filament current is increased, however, the beam current reaches a saturation value that is dependent on the geometry of the gun, and the value of the resistor R_c. Typical values of the components will be $R_c=0.5$ megohm, $C_c=0.5$ μf, for $r_k''=2.5$ mm. Under these conditions the source is definitely single and intense. Its position is nearly central with respect to the shield and is very little dependent on the centering of the filament with respect to the shield. The image seen at the intermediate screen is no longer the image of the filament but the image of a space charge.

An examination of the image at the intermediate screen of the microscope as the filament current is increased shows the course of events in the electron gun. At low filament current an image of the filament can be seen. It will generally show the die markings on the filament and will be in a position conjugate with the filament. As the filament current is increased this image becomes blurred, larger, and more intense, and moves to a new position, though no movement will be evident if the filament is carefully centered to the axis of the shield and anode holes.

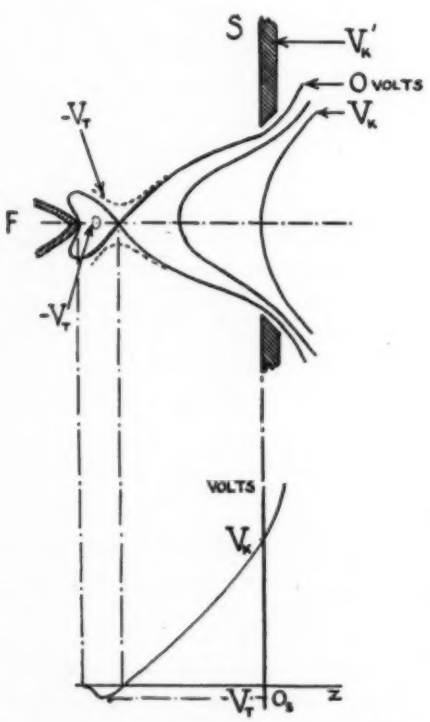


FIG. 16. Assumed potential distribution for a cathode-biased electron gun.

Compared with the zero-biased gun it is possible to obtain a brighter source even when the size of the source and the beam current is smaller. Increasing the shield-to-filament distance reduces the value at which the beam current saturates and also reduces the source size. Over a wide range of filament heights, however, the intensity of illumination is nearly constant, and it is both possible and advantageous to work with the filament raised so that the beam current is of the order 100 to 300 microamperes; the source radius is then of the order 5 to 15 microns and the intensity of the order twenty times that available with the zero-bias gun.

The potential distribution in this type of gun is shown diagrammatically in Fig. 16. Calculations of the potential distribution have been made and an empirical discussion of the behavior of the gun given by Reisner.[11] The space charge enclosed in the zero equipotential surface near the filament acts as a virtual cathode. Since electrons emitted with thermal velocity V_T (in electron volts) cannot cross the $-V_T$ equipotential surface this surface will tend to define the source. Moreover, since electrons cannot reach the shield no secondary sources can be produced by electron reflections there.

The presence of space charge and the ill-defined nature of the virtual cathode do not permit this type of gun to be discussed further in terms of the above theory.

THE EMISSION MICROSCOPE

Measurements of the performance of an immersion objective have been made by Johannson[12] and Savchenko.[13] In the emission microscope

[11] J. H. Reisner, Paper before the Electron Microscope Society of America, December 1946 meeting at Pittsburgh.
[12] H. Johannson, Ann. d. Physik **18**, 385 (1933).
[13] F. Savchenko, Zhurnal Tekhnicheskoi Fiziki, Moscow **9**, 2211–2219 (1939).

employed by Johannson an image of the emitting cathode was formed on a fluorescent screen 240 mm from the cathode. The constants for the microscope are listed below.

$d_2 = 1.0$ mm
$r_k'' = 0.6$ mm
$v_k = 240$ mm approximately $= L$ a constant for the microscope
$V = 750$ volts.

In operation u_k was set at some value between -5.0 mm and zero, and V_k' was adjusted to produce a sharp image on the fluorescent screen. Values of M_k, the magnification, and V_k' the screen potential for different values of u_k have been recorded by Johannson.

The foregoing theory must be extended before it can be applied to the above system. Since $v_k \gg d_2$ the electrons travel for the most part at constant speed after leaving the region of the shield. The transit time from the shield to the anode is small and will be ignored compared with the transit time from the anode to the fluorescent screen.

$$T_I = L/\{2d_2^{\frac{1}{2}}(-Ve/2md_2)^{\frac{1}{2}}\}, \tag{52}$$

and therefore

$$M_k = L[\lambda(1-\epsilon)]^{\frac{1}{2}}/2d_2^{\frac{1}{2}}u_k. \tag{53}$$

In Fig. 17 the points are calculated from Eq. (53), ϵ being deduced from the values of V_k' given by Johannson. The agreement between theory and experiment is good for $-u_k > r_k''$. With smaller values of u_k one may expect the assumptions made above as to the potential distribution to become increasingly invalid. In view of the large values of M_k, T_0 is much smaller than T_I, and the condition for focus is therefore approximately

$$T_0 = -T_F. \tag{54}$$

In calculating the focal time T_F it should be noticed that for this microscope V_k is not always small compared with V. Since the electron speed is constant on the image side of the lens, the focal time can be obtained in this case by calculating, first of all, the focal length F. If f_k, and f_A are the focal lengths of the cathode and anode components, respectively, then:

$$1/F = (1/f_k) + (1/f_A) \tag{55}$$

$$= \frac{(V-V_k)/d_2 + V_k/u_k}{4V_k} - \frac{(V-V_k)/d_2}{4V}. \tag{56}$$

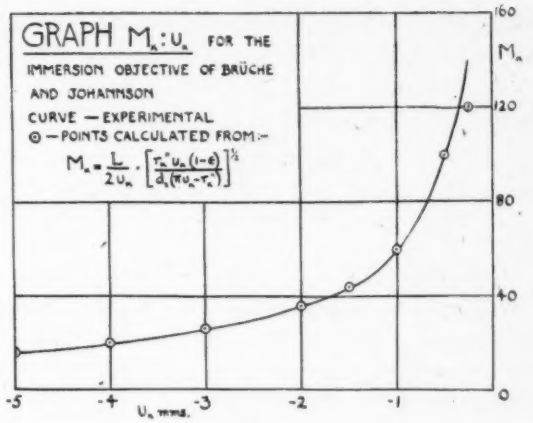


FIG. 17.

Also

$$T_F = F(m/-2eV)^{\frac{1}{2}} \tag{57}$$

and

$$V_k = r_k''u_k(1-\epsilon)/[\pi u_k - r_k''(1+u_k/d_2)]. \tag{58}$$

However Eq. (58) may be replaced by Eq. (13) with sufficient accuracy for the present calculations. With the substitutions previously employed

$$\frac{1}{F} = \frac{[1-\lambda(1-\epsilon)/d_2]^2u_k + \lambda(1-\epsilon)}{4u_k\lambda(1-\epsilon)}. \tag{59}$$

Solving for T_F and substituting into Eq. (54)

$$\frac{4\lambda^3(1-\epsilon)^3}{d_2^3} = \left\{\left[1-\frac{\lambda(1-\epsilon)}{d_2}\right]\frac{u_k}{d_2} + \frac{\lambda(1-\epsilon)}{d_2}\right\}^2. \tag{60}$$

Working with millimeters as the unit of length we have for the special case of the immersion

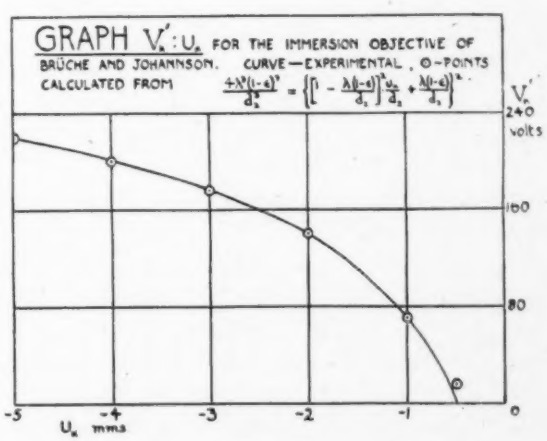


FIG. 18.

objective used by Johannson $d_2=1.0$. Hence writing $x=\lambda(1-\epsilon)$ Eq. (60) becomes

$$u_k^2x^4-(4u_k^2-2u_k+4)x^3 +(6u_k^2-4u_k+1)x^2 -(4u_k^2-2u_k)x+u_k^2=0. \quad (61)$$

Choosing values of u_k Eq. (61) is solved for x. From x and the known value of λ, ϵ and hence V_k' is deduced. Values of V_k' calculated in this way are plotted as points in Fig. 18—the line represents the values observed by Johannson. The agreement is good for $-u_k>r_k''$.

CONCLUSIONS

The chief conclusions reached with regard to the zero-bias electron gun used in electron microscopes are that no cross-over is formed in the gun under normal conditions of operation and that the small size of the effective source is due to the demagnification of the gun. The gun is inefficient since most of the electrons strike the condenser aperture. The dioptrics of such a system is not Gaussian. The cathode-biased gun is at present the most efficient type of gun for electron microscopy.

A satisfactory first-order theory of the immersion objective has been obtained for the case in which the distance from cathode to shield is greater than the radius of the shield aperture. The dioptrics of this system is also not Gaussian.

ACKNOWLEDGMENTS

It is a pleasure to record my indebtedness to Professor E. F. Burton, Director of the McLennan Laboratory for his encouragement during the prosecution of this work. Mr. R. S. Sennett, also of the McLennan Laboratory gave valuable help by making the electrolytic trough determinations. Finally I wish to thank Dr. James Hillier of the RCA Laboratories for useful information and informative discussions on the cathode-biased gun.

On the One-Dimensional Theory of Steady Compressible Fluid Flow in Ducts with Friction and Heat Addition*

BRUCE L. HICKS
Ballistic Research Laboratories, Aberdeen Proving Ground, Maryland
DONALD J. MONTGOMERY
Office of Naval Research (London Branch)
AND
ROBERT H. WASSERMAN
University of Chicago, Chicago, Illinois
(Received April 17, 1947)

Steady, diabatic (non-adiabatic), frictional, variable-area flow of a compressible fluid is treated in differential form on the basis of the one-dimensional approximation. The basic equations are first stated in terms of pressure, temperature, density, and velocity of the fluid. Considerable simplification and unification of the equations is then achieved by choosing the square of the local Mach number as one of the variables to describe the flow.

The transformed system of equations thus obtained is first examined with regard to the existence of a solution. It is shown that, in general, a solution exists whose calculation requires knowledge only of the variation with position of any three of the dependent variables of the system. The direction of change of the flow variables can be obtained directly from the transformed equations without integration. As examples of this application of the equations, the direction of change of the flow variables is determined for two special flows.

In the particular case when the local Mach number $M=1$, a special condition must be satisfied by the flow if a solution is to exist. This condition restricts the joint rate of variation of heating, friction, and area at $M=1$. Further analysis indicates that when a solution exists at this point it is not necessarily unique.

Finally, it is shown that the physical phenomenon of choking, which is known to occur in certain simple flow situations, is related to restrictions imposed on the variables by the form of the transformed equations. The phenomenon of choking is thus given a more general significance in that the transformed equations apply to a more general type of flow than has hitherto been treated.

INTRODUCTION

THE rational experimental development of jet- and rocket-propulsion power plants requires adequate knowledge of the theoretical mechanics of diabatic (non-adiabatic), frictional, variable-area compressible fluid flow. The differential equations describing this type of flow are well known.[1a, 1b, 2–4] Their solution in the three-dimensional case, however, is so difficult that some simplification is necessary to permit development of the theory in a form immediately useful for technical applications.

In the present paper, such simplification is effected by generalizing the familiar "one-dimensional" or hydraulic treatment of steady fluid flow to include the simultaneous effects of heat addition, friction, and area change upon the flow of a compressible fluid rather than by attempting to show that the one-dimensional approximation follows from a simplification of the hydrodynamic and heat-flow equations in their general three-dimensional form. The generalization leads to one-dimensional equations in differential form, which are identical with equations previously used by other investigators in less general cases.

Generalized conservation equations have been derived in Appendix A in order that a complete and logical basis for the theory may be accessible to the reader. The resulting theory is intended to serve as a foundation in differential form for calculation of all types of mathematically continuous (that is, shockless) flow of perfect gases to which the one-dimensional approximation is applicable. Thus the theory applies directly to

* This paper is a revised report of theoretical work performed by the authors at the Cleveland Laboratory of the National Advisory Committee for Aeronautics in 1944–45, and was issued in slightly different form as NACA TN No. 1336, 1947.

[1] W. F. Durand, ed. *Aerodynamic Theory* (Julius Springer Verlag, Berlin, 1935): (a) Vol. I, p. 237; (b) Vol. III, pp. 47–49.

[2] C. Eckart, Phys. Rev. **58**, 267 (1940).

[3] Wien-Harms, *Handbuch der Experimental Physik* (Akaddemische Verlagsgesellschaft, Leipzig, 1931), Bd. 4, pp. 343–369.

[4] A. Vazsonyi, Quart. Appl. Math. **3**, 29 (1945).

compressible flow in combustion chambers and also, with but slight modification, to flow in turbines, compressors,[5] nozzles, and diffusers, whenever the one-dimensional approximation is valid.

In order to obtain convenient and unified equations, the generalized relations are transformed by introducing a new basic variable, the square of the local Mach number $M^2=N$. Pressure and temperature are chosen as the additional basic variables; other relevant flow variables (for example, density, velocity, mass flow) may be expressed in terms of Mach number squared, pressure, and temperature. Values of M from zero to infinity are considered; the treatment is therefore applicable to both subsonic and supersonic flow.

The variable M has been used throughout differential treatments by Gukhman,[6] Bailey,[7] Nielsen,[8] and Chambre and Lin,[9] who investigated various examples of frictional, diabatic compressible flow.** A related variable

$$Z=(\gamma-1)M^2/[2+(\gamma-1)M^2],$$

which can be used alternatively with M, will be discussed briefly in Appendix B. Pertinent papers in which M is not used extensively are those by Binder[10] and Keenan and Neumann,[11] which report studies of isothermal and adiabatic frictional flow, respectively. A treatment of frictionless, diabatic compressible flow carried out by Szczeniowski[12] is partly in differential form. The same subject, in M language without differential formulation, is discussed by Hicks[13] and Wood.[14] The variable M has also been employed to advantage by Sorg[5] for the analysis of compressible flow through turbines and compressors, a related field that is not specifically discussed in the present paper. Mach and Crocco vectors, the square of whose magnitudes are N and Z, have been found useful for the description of three-dimensional adiabatic and diabatic flows.[15a, 15b]

In the general case, the differential equations obtained in the present treatment do not permit of formal integration; but being of first order, they are particularly amenable to numerical methods. A solution of the system is shown to exist, except possibly at sonic velocity, and the behavior of the solution in this neighborhood is investigated. From the differential equations useful information may easily be obtained about direction of changes in the flow variables. Choking is shown to be a consequence of a certain property of the equations.

THE ONE-DIMENSIONAL APPROXIMATION

Basic Equations

The "one-dimensional" steady-flow theory utilizes a model consisting of a perfect gas contained within a duct, across any section of which the flow variables are constant. Only the component of velocity normal to the section is considered; body forces are neglected, and heat, whether supplied by combustion, conversion of frictional work, or conduction from the walls, is assumed to be transferred instantaneously and completely but only transversely throughout the cross section, which may be of variable area. Each flow variable can thus be considered as a function of a single parameter, say the distance along the axis of the tube, whence the term "one-dimensional."

The conventional variables—pressure, temperature, density, and velocity in one-dimensional flow—are connected by four relations derivable from the first law of thermodynamics, the conservation of mass, the second law of motion, and the thermal equation of state for a perfect gas. The four relations are:

[5] K. W. Sorg, Forschung **10**, 270 (1939).
[6] A. A. Gukhman, J. Tech. Phys. U.S.S.R. **9**, 411 (1939).
[7] N. P. Bailey, J. Ae. Sci. **2**, 227 (1944).
[8] J. N. Nielsen, NACA ARR No. L4C16 (1944).
[9] P. Chambre and C. C. Lin, J. Ae. Sci. **13**, 537 (1946).
** The authors regret that explicit reference could not be made to several valuable (classified) papers by E. R. Hawthorne. It is our understanding that Dr. Hawthorne and Professor Shapiro have been invited to present a comprehensive paper on flow of compressible fluids with heat transfer at a meeting of the American Society of Mechanical Engineers later this year.
[10] R. C. Binder, A.S.M.E. Trans. **66**, 221 (1944).
[11] J. H. Keenan and E. P. Neumann, NACA Tech. Note No. 963 (1945).
[12] B. Szczeniowski, Can. J. Research **23**, 1 (1945).
[13] B. L. Hicks, NACA ACR No. E5A29 (1945) (Wartime Report No. E-88).
[14] G. P. Wood, J. Ae. Sci. **14**, 24 and 63 (1947).
[15] (a) B. L. Hicks, P. E. Guenther, and R. H. Wasserman, Quart. Appl. Math. (1947); (b) B. L. Hicks, Quart. Appl. Math., accepted for publication in 1948, see also Phys. Rev. **69**, 135 (A), 250 (A) (1946); **71**, 476 (A) (1947); Ballistic Research Laboratories Report No. 633 (1947).

Conservation of energy $c_p dT + VdV = dQ$, (1)

Conservation of mass $d(\rho VA) = 0$, (2)

Equation of motion $-dp = \rho VdV + \rho dF$, (3)

Equation of state $d(p/R\rho T) = 0$. (4)

The specific heat at constant pressure c_p and the gas constant R do not vary in the flow. The symbols ρ, V, p, and T, respectively, stand for density, velocity, absolute static pressure, and absolute static temperature. The pipe area, which may be variable, is represented by A. Heat added per unit mass is indicated by Q, and work per unit mass done against friction by F. Consistent units are used throughout. In Eqs. (1)–(4) each variable is to be considered as a function of a single parameter, such as the distance x along the tube considered positive in the direction of flow; and, of course, the meaning of each differential du is then given by

$$du = u'(x)dx.$$

Equations (1)–(3) are customarily used without explicit recognition of their true meaning with regard to the one-dimensional approximation. The interpretation of the quantity dF in particular is often obscure. In order to provide a logical, unified basis for the theory, Eqs. (1)–(3) are derived in Appendix A; special care is taken to keep the derivations within the framework of the one-dimensional approximation.

Applicability

The validity of the one-dimensional approximation depends upon the assumption of the uniformity of flow conditions across a plane normal to the direction of flow. Experience has shown that this assumption constitutes an adequate approximation in many special cases; in particular, with subsonic turbulent flow in long pipes without separation, the reasonably flat velocity profile permits the use of equations derived on this basis. Van Driest[16] has shown how the results of one-dimensional theory for incompressible fluids can be corrected for the effects of turbulence and non-uniformity of velocity distribution in cases where these factors have been evaluated experimentally. The corrections to the energy equation are quite complicated, and their complexity would be increased for compressible flow. In general there is insufficient experimental data available at present to permit formulation of such corrections where compressibility, incomplete growth of the boundary layer, or separation of the flow must be considered. Although compressibility and boundary layer effects are somewhat amenable to calculation, the occurrence of sparation is difficult or impossible to predict, and the question of applicability, in these cases, of the one-dimensional approximation must be determined by experiment or estimated by experience.

The one-dimensional approximation would not be valid if oblique shocks occur in the flow. Nor can normal shocks, if treated as flow discontinuities, be handled in the differential form of the present approximation. If, however, in Eqs. (1) and (3), dQ and dF are considered to depend upon the derivatives of T and V and if heat and momentum transfer in the direction of flow is allowed, then the equations for continuous normal shock[17] can be put in the form of Eqs. (1)–(4).

In the development and use of Eqs. (1)–(4) various approximations are made, such as neglecting the squares of velocity components normal to the direction of flow, replacing the square of the cosine of the half-angle by unity, and assuming the constancy of R and c_p. In this paper no attempt is made to state under what circumstances such approximations are suitable.

TRANSFORMATION OF EQUATIONS

Change of Variables

A canonical form for Eqs. (1)–(4) is obtained by taking logarithmic derivatives and choosing as a variable the square of the local Mach number,

$$M^2 \equiv N = V^2/\gamma RT, \quad (5)$$

where γ is the ratio of the specific heats. This choice to obtain simplification of the equations is not unique; similar advantages result with other dimensionless combinations of velocity squared and a temperature. For instance, some workers have used the ratio of dynamic tem-

[16] E. R. Van Driest, A.S.M.E. Trans. **68A**, 231 (1946).

[17] Reference 1, p. 219.

perature to total temperature; in Appendix B to the present report, the canonical differential equations in terms of this variable are given.

If Eqs. (1)–(4) are divided by c_pT, ρVA, p, and $p/R\rho T$, respectively, there result

$$(\gamma-1)\frac{V^2}{\gamma RT}\frac{dV}{V}+\frac{dT}{T}=\frac{dQ}{c_pT}, \tag{6}$$

$$\frac{dV}{V}+\frac{d\rho}{\rho}=-\frac{dA}{A}, \tag{7}$$

$$\gamma\frac{V^2}{\gamma RT}\frac{dV}{V}+\frac{dp}{p}=-\frac{\rho dF}{p}, \tag{8}$$

$$-\frac{dp}{p}+\frac{dT}{T}+\frac{d\rho}{\rho}=0. \tag{9}$$

With use of Eq. (5) and the expression for dV/V obtained by logarithmic differentiation of Eq. (5),

$$\frac{dV}{V}=\frac{1}{2}\left(\frac{dN}{N}+\frac{dT}{T}\right) \tag{10}$$

and, upon elimination of $d\rho/\rho$, there are found

$$\frac{(\gamma-1)N}{2}\frac{dN}{N}+\left[1+\frac{(\gamma-1)N}{2}\right]\frac{dT}{T}=\frac{dQ}{c_pT}\equiv d\theta, \tag{11}$$

$$\frac{1}{2}\frac{dN}{N}+\frac{dp}{p}-\frac{1}{2}\frac{dT}{T}=-\frac{dA}{A}\equiv d\alpha, \tag{12}$$

$$-\frac{\gamma N}{2}\frac{dN}{N}-\frac{dp}{p}-\frac{\gamma N}{2}\frac{dT}{T}=\frac{dF}{RT}\equiv d\mu, \tag{13}$$

where the dimensionless quantities $d\theta$, $d\alpha$, and $d\mu$ have been introduced to simplify the following analysis.

Solution for Logarithmic Differentials

If the determinant formed by the coefficients of dN/N, dp/p, and dT/T in Eqs. (11), (12), and (13) is not identically zero, the equations may be solved uniquely for these three differentials. As the determinant in question is proportional to $(1-N)$, which vanishes only for $N=1$, the solution is obtained as follows:

$$dN/N=(1-N)^{-1}\{(1+\gamma N)d\theta+[2+(\gamma-1)N]d\mu+[2+(\gamma-1)N]d\alpha\}, \tag{14}$$

$$dp/p=(1-N)^{-1}\{-\gamma Nd\theta-[1+(\gamma-1)N]d\mu-\gamma Nd\alpha\}, \tag{15}$$

$$dT/T=(1-N)^{-1}[(1-\gamma N)d\theta-(\gamma-1)Nd\mu-(\gamma-1)Nd\alpha]. \tag{16}$$

It is also convenient to record the differential expressions for the density ρ and velocity V:

$$d\rho/\rho=dp/p-dT/T=(1-N)^{-1}\times(-d\theta-d\mu-Nd\alpha), \tag{17}$$

$$dV/V=(dN/N+dT/T)/2=(1-N)^{-1}(d\theta+d\mu+d\alpha). \tag{18}$$

Application of Second Law of Thermodynamics

The first law of thermodynamics was used in the formulation of the basic equations; the second law of thermodynamics may be employed to furnish additional information. The entropy differential dS for a perfect gas is given[18] by

$$dS/c_p=dT/T-[(\gamma-1)/\gamma]dp/p=d\theta+[(\gamma-1)/\gamma]d\mu. \tag{19}$$

The second law of thermodynamics then states

$$0\leqslant dS/c_p-dQ/c_pT=[(\gamma-1)/\gamma]d\mu. \tag{20}$$

The relation, according to Eq. (19), that

$$dS/c_p=(d\theta+d\mu+d\alpha)-d\mu/\gamma-d\alpha$$

when used with Eq. (20), results in the inequalities

$$d\theta\leqslant dS/c_p=d\beta-d\mu/\gamma-d\alpha\leqslant d\beta-d\alpha, \tag{21}$$

where $d\beta\equiv d\theta+d\mu+d\alpha$.

DISCUSSION OF EQUATIONS

Remarks on Integration of Equations

Equations (14)–(16) can be rewritten as

$$N'=\frac{N}{1-N}\frac{1+\gamma N}{c_pT}Q'+\frac{N}{1-N}\frac{2+(\gamma-1)N}{RT}F'-\frac{N}{1-N}\frac{2+(\gamma-1)N}{A}A', \tag{22}$$

[18] P. S. Epstein, *Textbook of Thermodynamics* (John Wiley and Sons, Inc., New York, 1937), p. 63.

$$p' = -\frac{p}{1-N}\frac{\gamma N}{c_p T}Q' - \frac{p}{1-N}\frac{1+(\gamma-1)N}{RT}F' + \frac{p}{1-N}\frac{\gamma N}{A}A', \quad (23)$$

$$T' = \frac{T}{1-N}\frac{1-\gamma N}{c_p T}Q' - \frac{T}{1-N}\frac{(\gamma-1)N}{RT}F' + \frac{T}{1-N}\frac{(\gamma-1)N}{A}A', \quad (24)$$

where the primes indicate differentiation with respect to x. This system clearly satisfies, except at $N=1$, the conditions of the fundamental existence theorem[19] when Q, F, and A are differentiable. Hence a solution exists except at sonic velocity and may be obtained formally when possible, and by standard numerical methods otherwise, as soon as the functions Q, F, and A (or their derivatives) are specified. More generally, the system may be solved in similar fashion for any three of the variables N, p, T, Q, F, and A as functions of x, when the variation with x of the other three is prescribed. Also it may be noted that as all the foregoing variables are functions of one parameter, any two may be considered as functions of each other under suitable circumstances.

Direction of Change of Flow Variables

In practical as well as in theoretical work it is frequently useful to be able to determine the direction of change of flow quantities with respect to heat addition, friction, or area variation without troubling to get quantitative information from integrated forms. Equations (14)–(18) or (22)–(24) permit the specification of signs of derivatives at any particular point and also throughout certain regions of flow. Thus Eq. (14) shows that in subsonic flow the effect of positive θ', μ', or α', is to increase N, whereas for supersonic flow the effect is to decrease N. When the derivatives have different signs, the net effect will depend upon the algebraic sum of the separate contributions.

As an example of the use of this technique, suppose heat is added to a fluid in a constant-area pipe, with negligible friction; that is, $\theta' \not\equiv 0$, $\mu' = \alpha' \equiv 0$. It is easily seen from Eqs. (14) through (18) that for the entire range of N from zero to infinity

$$(1-N)dN/dQ \gtreqless 0, \quad (25)$$

$$dp/dN \lesseqgtr 0, \quad (26)$$

$$(1-\gamma N)^{-1}dT/dN \gtreqless 0, \quad (27)$$

$$d\rho/dN \lesseqgtr 0, \quad (28)$$

$$dV/dN \gtreqless 0. \quad (29)$$

These results are given by Hicks.[13] By use of the chain rule for the derivative of a function of a function, the sign of the derivative of any of the flow variables with respect to any of the others may be obtained; thus, from Eqs. (25) and (29) it is clear that

$$(1-N)\frac{dV}{dQ} = (1-N)\frac{dV}{dN}\frac{dN}{dQ} \gtreqless 0. \quad (30)$$

As another example, consider the flow in circular cylindrical pipes with heat addition and with friction; that is, $\theta' \not\equiv 0$, $\mu' \not\equiv 0$, $\alpha' \equiv 0$. (See also related discussion by Nielsen.)[8] Equation (14) will be used to determine the direction of change of N with respect to x. If the heat addition is only through the wall, which is at temperature T_w, the heat added per unit mass of fluid in passing a distance dx along the tube is given by

$$\rho V A dQ = h(T_w - T)(\pi D dx), \quad (31)$$

where D is the tube diameter, and h the local surface-to-fluid coefficient of heat transfer in heat units transferred per unit temperature difference, per unit area.[20a] In conjunction with Eq. (11), Eq. (31) leads to

$$d\theta \equiv \frac{dQ}{c_p T} = \frac{h[(T_w/T)-1]\pi D dx}{c_p \rho V \pi D^2/4} = \frac{4h[(T_w/T)-1]dx}{c_p \rho V D}. \quad (32)$$

The expression for frictional work done is

[19] L. Bieberbach, *Theorie der Differentialgleichungen* (Julius Springer Verlag, Berlin, 1930), pp. 34–35.

[20] W. H. McAdams, *Heat Transmission* (McGraw-Hill Book Company, Inc., New York, 1942). (a) Eq. 2, p. 135; (b) Eq. 8, p. 119; (c) Eq. 1, p. 162.

assumed to be given[20b] (see also Appendix A) by

$$dF=\frac{fV^2dx}{2(D/4)}, \tag{33}$$

where f is the Fanning friction factor. From Eqs. (13) and (33) it follows that

$$d\mu\equiv\frac{dF}{RT}=\frac{2f\gamma Ndx}{D}. \tag{34}$$

If Reynolds' analogy is valid, h may be replaced[20c] by $c_p\rho Vf/2$, whence Eq. (14) becomes

$$dN/dx=[N/(1-N)]\{(1+\gamma N)[(T_w/T)-1] +[2+(\gamma-1)N]\gamma N\}2f/D. \tag{35}$$

This equation may be used to determine the direction of change of N with x, and hence of other flow quantities, for various ranges of N and of T_w/T. For values of $(T_w/T)\ll 1$, (maximum rate of cooling), dN/dx is positive for values of

$$1>N>[-\gamma+(\gamma(5\gamma-4))^{\frac{1}{2}}]/2\gamma(\gamma-1)=0.58$$

for $\gamma=1.4$; that is, the effects of friction in increasing the Mach number overbalance the effects of the cold walls in lowering it if $1>M=N^{\frac{1}{2}}>0.76$ for $\gamma=1.4$ If $N>1$ then $(dN/dx)<0$, and acceleration of frictional, supersonic flow by convective cooling appears to be impossible. This behavior of compressible gases affects the design of supersonic wind tunnels.[21] Acceleration of frictionless supersonic flow by cooling should, however, be possible[9,13].

Behavior of Solution at Sonic Velocity

The differential equations (14)–(18) must be examined as to behavior at the singular point $N=1$. In order that the logarithmic differentials may be defined at this point, it is necessary that

$$d\beta\equiv d\theta+d\mu+d\alpha=0 \text{ at } N=1, \tag{36}$$

because each logarithmic differential is proportional to $d\beta/(1-N)$ there. If $d\beta\neq 0$ upstream of the end of the duct, N can equal 1 only at the end of the duct. This situation is illustrated by the "choked" converging nozzle and by the frictional diabatic flow, which is treated in the previous section. Equation (36) is formally satisfied at the end of a duct where $d\theta$, $d\mu$, and $d\alpha$ may be considered to vanish for all values of N.

Between the ends of a duct, however, $d\beta$ must always vanish where $N=1$. This condition shows that at $N=1$ arbitrary variations of $d\theta$, $d\mu$, and $d\alpha$ are not possible. A specific illustration is the ideal nozzle in which $d\theta=d\mu=0$; according to Eq. (36), $d\alpha$ is then restricted to the value 0, which means that the area has a stationary value at $N=1$. This is the well-known result that sonic velocity can be attained only in the throat of an ideal nozzle in shockless flow. A quite similar treatment applies for the cases where $d\theta$ and $d\mu$ are the quantities to be investigated (see pertinent material in references 5–14 and 21). Condition (36), which was necessitated by the presence of the determinant $(1-N)/2$ in Eqs. (14)–(18), is thus seen to provide a unification of the treatment of the flow behavior in the neighborhood of sonic velocity.

Combination of the second law of thermodynamics with Eq. (36) also yields limitations on the behavior of the flow at $N=1$. According to Eqs. (21) and (36), at sonic velocity

$$d\theta\leqq dS/c_p\leqq -d\alpha=dA/A. \tag{37}$$

These results may be stated in words to the effect that in converging or constant-area channels at $N=1$, neither the heat term $d\theta=dQ/c_pT$ nor the entropy term dS/c_p can be positive. In diverging channels these two terms may be either positive or negative. If either $d\theta$ or $d\mu$ is everywhere 0, relation (21) yields more detailed results. For example, if $d\theta\equiv 0$, then at $N=1$, by Eqs. (36) and (20)

$$d\mu=-d\alpha =[\gamma/(\gamma-1)]dS/c_p\geqq 0, \quad (d\theta=0,\ N=1).$$

Continuous flow with friction and without heat addition at sonic velocity cannot, therefore, occur in a converging channel.[7]

A more complete treatment of the behavior of the flow when $N=1$ between the ends of the duct is obtained by considering second-order terms. As N approaches unity, Eq. (14), which may be written

$$N'/N=\{(1+\gamma N)\theta'+[2+(\gamma-1)N]\mu' +[2+(\gamma-1)N]\alpha'\}/(1-N),$$

[21] F. Clauser, Phys. Rev. 71, 465A (1947).

takes the form 0/0. For the evaluation of this limit, L'Hopital's rule gives after some calculation

$$N_0' = [\theta_0' N_0' + (1+\gamma)\beta_0'']/(-N_0'),$$

where subscript 0 denotes value of function at $N=1$. The solution for N_0' is

$$N_0' = -\theta_0'/2 \pm [(\theta'_0/2)^2 - (1+\gamma)\beta_0'']^{\frac{1}{2}}. \quad (38)$$

The double-valuedness of the derivative at $N=1$ will have important consequences in that a unique solution of the equations may not be obtained when $N=1$ along the flow path. In general, it will be possible to continue the solution from $N=1$ along either of two paths, depending on the choice of sign. In certain cases, depending on the signs of θ_0' and β_0'', one sign will correspond to continuation into subsonic flow, the other into supersonic; otherwise the two choices will correspond to different continuations into flow of the same character. This result means that specification of initial conditions and of variation of $d\theta$, $d\mu$, and $d\alpha$ alone is not sufficient to insure a unique solution if N becomes unity along the flow. In the event that the radicand is zero, it is possible that only one solution is obtained; or it may happen that some higher derivative is double-valued with resulting ambiguity of solution. The analysis for this case is somewhat involved and will not be continued here.

It is interesting to note that a less general problem of the same nature has been presented by Lorenz[22] and Prandtl and Proell.[23] Some of this work is possibly more accessible in Stodola's book.[24]

The Phenomenon of Choking

The general equations (11)–(13) impose restrictions on the relations between the flow variables and the heat, friction, and area variation. When these restrictions take the form of upper or lower limits on mass flow, the associated phenomena are termed "choking" processes. As an example, it is well known that the ideal nozzle has for given subsonic entry conditions a maximum mass flow beyond which the discharge cannot be increased no matter how much the exit pressure is lowered. Another case is "thermal choking," wherein the entrance Mach number and hence mass flow in diabatic, frictionless, constant-area flow is limited for given heat addition despite indefinite reductions in outlet pressure.[13]

The nature of choking may be studied with the help of Eq. (22), which was derived simply from the basic equations. It will be shown that unless heat, friction, and area variation are such that $(1-N)$ times the right-hand side of Eq. (22) changes from positive to negative as x increases, the Mach number in the tube cannot become greater than 1 if the entrance velocity is subsonic and cannot become less than 1 if the entrance velocity is supersonic, provided that the flow variables remain continuous.

For convenience, designate by Y the factor $(1-N)$ times the right-hand side of Eq. (22). The quantity Y is seen to consist of a sum of terms in Q', F', and $-A'$ multiplied by functions of N that are always positive. (In the event that only one of the terms Q', F', and $-A'$ is not 0, Y becomes merely the derivative of the heat added, the frictional work, or the area, multiplied by a simple function of the flow variables; then positive Y corresponds to the case of heat addition, friction, or a converging duct.)

Suppose first that Y is always negative. Then if the flow at the entrance section x_1 is subsonic, $dN/dx = (1-N)^{-1}Y < 0$, and the Mach number decreases; if the entering flow is supersonic, $dN/dx = (1-N)^{-1}Y > 0$, and the Mach number increases.

Suppose now that Y is always positive. Then, if the entering flow at x_1 is subsonic, $dN/dx = (1-N)^{-1}Y > 0$, and the Mach number increases. But N cannot increase past unity as x increases. For suppose $N=1$ at $x=x_0$ and is greater than 1 in the right-hand neighborhood of x_0 (exclusive of x_0); then dN/dx is negative in this neighborhood, because $(1-N)$ is less than 0 and Y is greater than 0. Now N is equal to unity at $x=x_0$, is continuous, and has a negative derivative in the neighborhood mentioned. Hence N is less than 1 in this neighborhood, which contradicts the assumption. Therefore N cannot be

[22] H. Lorenz, Physik. Zeits. **4**, 333 (1903).
[23] L. Prandtl and R. Proell, V.D.I. Zeits. **48**, 348 (1904).
[24] A. Stodola, *Steam and Gas Turbines* (McGraw-Hill Book Company, Inc., New York, 1927), pp. 98–101.

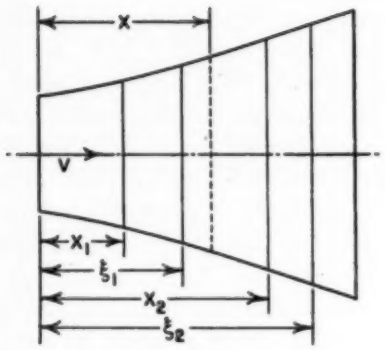


FIG. 1. Fluid elements considered in deriving the equations of energy and mass conservation.

greater than 1 if Y is always positive, N is continuous, and $N(x_1)$ is less than 1. In general, no continuous solution exists for values of $x>x_0$ if Y is always positive. This statement, and the foregoing proof, are valid even if $N'(x_0)$ does not exist. An analogous development may be made for $N(x_1)$ greater than 1 with the conclusion that, with Y positive and N continuous, N cannot be less than 1.

If Y changes from positive to negative at $x=x_0$, however, the value of N may cross unity at that point, but if Y is initially negative, N goes away from unity as previously shown and can only turn toward unity if Y changes from negative to positive; this change must be made at some value of N other than 1. After Y has changed to positive, the situation reduces to the case that Y is always positive, with the N where Y changes sign now taken as the entrance N.

It has been shown that up to some fixed point in the tube, which can be either the exit or the point at which Y changes from positive to negative, N and, therefore, the Mach number do not become greater than 1 if the entering velocity is subsonic nor less than 1 if the entering velocity is supersonic. Furthermore, if Y is positive up to the point at which N is limited, the derivative of N before this point is always positive if the entering flow is subsonic and is always negative if the entering flow is supersonic. Thus, for positive Y and subsonic entrance velocity the entrance N cannot exceed some limiting value less than 1 determined by the particular Q, F, $-A$ variation, for N is always increasing from its initial value and cannot exceed 1; and, by analogous considerations for positive Y and supersonic entrance velocity N cannot be less than some limiting value greater than 1. This limitation is essentially the choking phenomenon.

The specific form this limitation takes is not easily stated in the general case, because the choice of which factors are to be held constant and of which variables are to be considered limited determines the particular form of the restrictions. Numerical results for some particular cases are given in references 7, 13, and 14, among others, to illustrate the nature of possible results. It is felt that additional special cases should be investigated before a thorough study of the general case is attempted.

APPENDIX A

Derivations of Three Basic Equations

Conservation of Energy

The first law of thermodynamics applies to energy changes between two states of a system enclosed within a surface. Let the system (Fig. 1) be the gas of mass Δm that is contained in the initial state within the tube walls and the sections at x_1 and x_2 and in the final state within the tube walls and the sections at ξ_1 and ξ_2; x_1, x_2, and ξ_1 are arbitrary, and ξ_2 is determined by the condition that the mass between x_1 and ξ_1 equal the mass between x_2 and ξ_2. When $m(x)$ is defined as the total mass of gas contained within the duct between the sections $x=0$ and $x=x$, the definition of Δm becomes

$$\Delta m \equiv m(x_2)-m(x_1)=m(\xi_2)-m(\xi_1). \quad (39)$$

Let $U(x)$ denote the internal (thermal) energy per unit mass of fluid, $p(x)$ the static pressure, $V(x)$ the gas velocity, $A(x)$ the cross-sectional area of the duct, each taken at the section x; and $Q(x)$ the heat (in mechanical-energy units) added from walls or by combustion*** to a unit mass of the fluid during its passage from $x=0$ to $x=x$. Then the first law of thermodynamics says for

*** Actually, the heat liberated by combustion might be considered as part of the internal energy; or the external surroundings might be considered to include the fuel; or the first law might be generalized to include heat sources. The treatment given here is convenient but must be understood to require some justification on one of the bases mentioned.

steady flow that

$$\int_{m(\xi_1)}^{m(\xi_2)}(U+\tfrac{1}{2}V^2)dm-\int_{m(x_1)}^{m(x_2)}(U+\tfrac{1}{2}V^2)dm$$
$$=\int_{m(\xi_1)}^{m(\xi_2)}Qdm-\int_{m(x_1)}^{m(x_2)}Qdm$$
$$-\left[\int_{x_2}^{\xi_2}pAdx-\int_{x_1}^{\xi_1}pAdx\right]. \quad (40)$$

As $Adx=(1/\rho)dm$, where $\rho(x)$ is the density of the gas, Eq. (40) becomes

$$\int_{m(x_2)}^{m(\xi_2)}(U+\tfrac{1}{2}V^2-Q+p/\rho)dm$$
$$=\int_{m(x_1)}^{m(\xi_1)}(U+\tfrac{1}{2}V^2-Q+p/\rho)dm \quad (41)$$

when the limits of integration are changed.

Provided that the integrand is continuous (which requirement excludes shock), the expression on the right-hand side of Eq. (41) may be written, by the theorem of the mean for integrals:

$$[m(\xi_1)-m(x_1)]f(m_1^*),$$

where $f(m_1^*)$ indicates the value of the integrand at some $m=m_1^*$, $m(x_1)<m_1^*<m(\xi_1)$. The integral on the left-hand side yields a similar result, with subscript 1 replaced by subscript 2, whence, by virtue of Eq. (39)

$$f(m_1^*)=f(m_2^*);\quad m(x_1)<m_1^*<m(\xi_1),$$
$$m(x_2)<m_2^*<m(\xi_2). \quad (42)$$

As $\xi_1\to x_1$, $\xi_2\to x_2$, this equation becomes

$$f[m(x_1)]=f[m(x_2)]. \quad (43)$$

But x_1 and x_2 are arbitrary. Hence f is a constant, and

$$\frac{d}{dx}(U+\tfrac{1}{2}V^2-Q+p/\rho)=0$$

or, since $d(U+p/\rho)=dH$, where H is the enthalpy per unit mass,

$$dQ=dH+VdV. \quad (44)$$

For a perfect gas, $dH=c_p dT$, whence the energy equation is finally

$$dQ=c_p dT+VdV. \quad (45)$$

Conservation of Mass

The conservation of mass, in the form useful here, states that in a steady flow the mass entering a closed surface during any time interval Δt is equal to the mass leaving during the interval Δt. Let the closed surface consist of the sections at arbitrary ξ_1 and x_2 (Fig. 1) and the portions of the duct between these sections, and let Δt be the time required for the mass $m(\xi_1)-m(x_1)$ to enter the surface while mass $m(\xi_2)-m(x_2)$ flows out. The value of x_1 is arbitrary, whereas ξ_2 is fixed by the conditions on the time interval. Upon definition of $t(x)$ as the time required for a fluid particle to travel from origin $x=0$ to $x=x$, Δt may be defined by

$$\Delta t\equiv t(\xi_1)-t(x_1)=t(\xi_2)-t(x_2). \quad (46)$$

The law of conservation of mass says that

$$\int_{x_1}^{\xi_1}\rho Adx=\int_{x_2}^{\xi_2}\rho Adx. \quad (47)$$

Upon change of variable, this equation becomes

$$\int_{t(x_1)}^{t(\xi_1)}\rho A\frac{dx}{dt}dt=\int_{t(x_2)}^{t(\xi_2)}\rho A\frac{dx}{dt}dt. \quad (48)$$

As before, by the theorem of the mean for integrals and condition (46), the integrand must be constant, whence

$$d(\rho VA)=0. \quad (49)$$

Equation of Motion

The vector form of the second law of motion for continuous media states that the integral of the density of surface forces over a closed surface

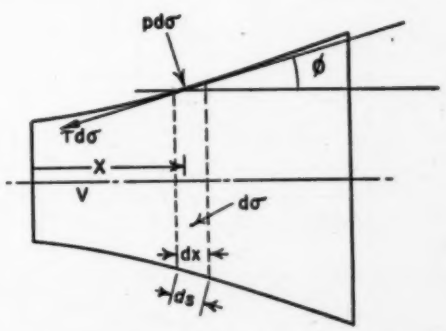


FIG. 2. Forces acting upon a fluid element described in the one-dimensional approximation.

is equal to the integral of the density times the particle derivative of the velocity over the volume enclosed by the surface.[25] For the mass of gas contained within the sections at x_1 and x_2 (Fig. 1) and the walls of the duct, the horizontal component of the equation of motion becomes

$$p(x_1)A(x_1)-p(x_2)A(x_2)+\int_{x_1}^{x_2} Rdx = \int_{x_1}^{x_2} \rho\frac{dV}{dt}Adx, \quad (50)$$

where $R(x)dx$ is the horizontal component of the force exerted on the gas by the portion of the duct between x and $x+dx$. For steady flow, the integral on the right-hand side may be transformed as follows:

$$\int_{x_1}^{x_2} \rho\frac{dV}{dt}Adx=\int_{x_1}^{x_2} \rho\frac{dV}{dx}\frac{dx}{dt},$$

$$Adx=\rho VA\int_{x_1}^{x_2}\frac{dV}{dx}dx=\rho VA[V(x_2)-V(x_1)]. \quad (51)$$

It follows from Eqs. (50) and (51) that the equation of motion may be written in differential form as

$$pdA+Adp+\rho VAdV=Rdx. \quad (52)$$

Now Rdx may be resolved into two constituent parts, the horizontal components of the tangential frictional drag and of the normal pressure reaction. If the half-angle of the duct is denoted by $\phi(x)$, the wall surface by $\sigma(x)$, and the tangential frictional drag per unit area by $\tau(x)$, from Fig. 2 it is clear that

$$Rdx=-(\tau d\sigma)\cos\phi+(pd\sigma)\sin\phi,$$

and that

$$d\sigma=dA/\sin\phi,$$

hence

$$Rdx=-(\tau d\sigma)\cos\phi+pdA. \quad (53)$$

It is possible to use Eq. (53) directly without further analysis if the friction factor is related to τ by the equation $\tau=fpV^2/2$. In many engineering treatments, however, f is defined in terms of the "energy loss due to friction." In order to make this concept of energy loss more precise and to make possible derivation of a rigorous connection between τ and energy loss, define $F(x)$ as the work done by unit mass of the fluid against friction in moving from the origin to position x. The work done in moving the entire mass of fluid between x_1 and ξ_1 to the region between x_2 and ξ_2 (see Fig. 3) will be computed in terms of the original variables τ and σ and in terms of the new variables F and m. If the two quantities are equated and suitably transformed, a relation will be obtained between dF and $\tau d\sigma$.

Let x_1, ξ_1, and x_2 be chosen arbitrarily, and let ξ_2 be determined by the condition that the total mass Δm between x_1 and ξ_1 equal the total mass between x_2 and ξ_2. Let x_a be an arbitrary point between x_1 and ξ_1, and let x_b be determined by the condition that $m(x_a)-m(x_1)=m(x_b)-m(x_2)$. In particular, if $x_a=x_1$, then $x_b=x_2$; if $x_a=\xi_1$, then $x_b=\xi_2$. Thus it is seen that, as x_a runs from x_1 to ξ_1, x_b runs from x_2 to ξ_2.

In order to determine the work done in terms of the variables τ and σ, the procedure is to move thin sections from their original positions, given in each case by x_a, to their final positions x_b, to find the work done by each of them, and then to add up the work done for all the sections.

First, divide the mass Δm between x_1 and ξ_1 into smaller elements of mass $\Delta_1 m$, $\Delta_2 m$, $\cdots\Delta_n m$. (See Fig. 3.) Let $\Delta_i\sigma$ be the wall surface corresponding to the element of mass $\Delta_i m$, and let the duct wall between the initial and the final position of $\Delta_i m$ be divided into elements of length $\Delta_j s$. For a given $\Delta_i m$ it follows from Fig. 3

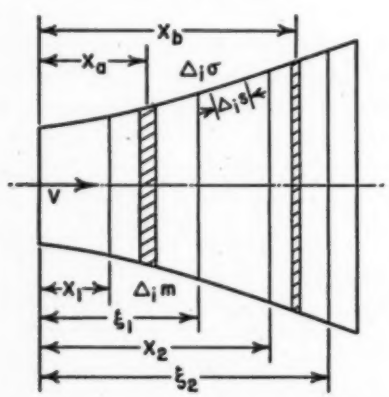


FIG. 3. Fluid elements considered in deriving the equation of momentum conservation.

[25] H. V. Craig, *Vector and Tensor Analysis* (McGraw-Hill Book Company, Inc., New York, 1943), pp. 355–356.

that

$$\Delta_i m = \rho_i A_i \Delta_i x,$$
$$\Delta_i \sigma = 2\pi r_i \Delta_i s,$$
$$\Delta_i x = \cos\phi_i \Delta_i s,$$

so that

$$\Delta_i \sigma = g_i \Delta_i m,$$

where $g_i \equiv 2\pi r_i / \rho_i A_i \cos\phi_i$. Each of the quantities ρ_i, r_i, ϕ_i, and g_i is to be evaluated at the proper point within $\Delta_i x$.

Finally, from Fig. 3 there exists on $\Delta_i m$ a force $\Delta F_i(s_j)$, to be overcome by the work against friction, equal to

$$\tau(s_j)\Delta_i\sigma(s_j) = \tau(s_j) g_i(s_j) \Delta_i m,$$

where it is convenient to consider τ and σ as functions of s, inasmuch as frictional drag and the extension of an element of fixed mass depends on the position of the element in the tubes.

For any element of mass $\Delta_i m$ then, an approximation (the accuracy of which depends on the size of $\Delta_i m$) to the work done by $\Delta_i m$ in going from section x_a to section x_b is

$$\lim_{\Delta_j s \to 0} \sum_{s(x_a)}^{s(x_b)} \Delta F_i(s_j) \Delta_j s = \lim_{\substack{k \to \infty \\ \Delta_j s \to 0}} \sum_{j=1}^{k} \tau(s_j) g_i(s_j) \Delta_i m \Delta_j s$$
$$= \Delta_i m \int_{s[m(x_{a_i})]}^{s[m(x_{b_i})]} \tau(s) g_i(s) ds.$$

It follows that the corresponding approximation to the work done by the entire mass Δm is

$$\sum_{i=1}^{n} \left\{ \int_{s[m(x_{a_i})]}^{s[m(x_{b_i})]} \tau(s) g_i(s) ds \right\} \Delta_i m$$

and, as the approximation becomes more and more accurate as the largest $\Delta_i m$ becomes smaller and smaller, the expression for the work done against friction when the entire mass between x_1 and ξ_1 is transported to the position between x_2 and ξ_2 becomes

$$\lim_{\substack{n \to \infty \\ \Delta_i m \to 0}} \sum_{i=1}^{n} \left\{ \int_{s[m(x_{a_i})]}^{s[m(x_{b_i})]} \tau(s) g_i(s) ds \right\} \Delta_i m$$
$$= \int_{m(x_1)}^{m(\xi_1)} \int_{s[m(x_a)]}^{s[m(x_b)]} \tau(s) g(s) ds dm. \quad (54)$$

It is clear that the work done against friction when the entire mass is moved as previously described is equal also to

$$\lim_{\substack{\Delta_j F \to 0 \\ \Delta_i m \to 0}} \sum_i \sum_j \Delta_j F \Delta_i m = \int_{m(x_1)}^{m(\xi_1)} \int_{F[m(x_a)]}^{F[m(x_b)]} dF dm$$
$$= \int_{m(x_1)}^{m(\xi_1)} \int_{F[m(x_a)]}^{F[m(x_b)]} (dF/ds) ds dm \quad (55)$$

whence it follows, since the intervals for both integrations are arbitrary, that

$$dF/ds = \tau g \quad (56)$$

and

$$dF = \frac{2\pi r}{\rho A \cos\phi} \frac{\tau d\sigma}{2\pi r}$$

whence

$$(\tau d\sigma) \cos\phi = \cos^2\phi \rho A dF. \quad (57)$$

Because, for the small angles usually under consideration, $\cos^2\phi$ is very nearly unity (for a half-angle of 6°, $\cos^2\phi = 0.989$), the retention of this factor except for particularly precise work would not seem justifiable. Hence Eq. (53) may be written

$$R dx = -\rho A dF + p dA. \quad (58)$$

The differential equation of motion is finally

$$-dp = \rho V dV + \rho dF. \quad (59)$$

The connection between dF and the differential loss in stagnation pressure $(-dp_t)$ can, with the help of Eqs. (14) and (15), be expressed in the form

$$\frac{-dp_t}{p_t} = \frac{dF}{RT} + \frac{\gamma N d\theta}{[2+(\gamma-1)N]}. \quad (60)$$

Thus except for the limiting case of incompressible flow, $(-dp_t)$ and dF cannot be used interchangeably in defining the friction factor even for the adiabatic case where $d\theta = 0$.

APPENDIX B

The Z Language

In the place of $N = V^2/\gamma RT$ the equations may be formulated in terms of Z, defined through

$$Z = (V^2/2c_p)/(T + V^2/2c_p). \quad (61)$$

The numerator is the so-called dynamic temperature, the denominator is the total tem-

perature. Z and N are related by

$$Z=[(\gamma-1)N]/[2+(\gamma-1)N], \quad (62)$$

$$N=2Z/(\gamma-1)(1-Z). \quad (63)$$

This replacement represents a one-to-one transformation of N into Z in the range 0 to infinity for N, 0 to 1 for Z.

In order to illustrate the form that some of the earlier equations take under this transformation, Eqs. (14)–(16) are written in terms of Z,

$$\frac{dZ}{Z}=\frac{2}{1-[(\gamma+1)/(\gamma-1)]Z}\left\{\frac{1-Z}{2}\left[1+\left(\frac{\gamma+1}{\gamma-1}\right)Z\right]d\theta+(1-Z)d\mu+(1-Z)d\alpha\right\}, \quad (64)$$

$$\frac{dp}{p}=\frac{2}{1-[(\gamma+1)/(\gamma-1)]Z}\left(-\frac{\gamma}{\gamma-1}Zd\theta-\frac{1+Z}{2}d\mu-\frac{\gamma}{\gamma-1}Zd\alpha\right), \quad (65)$$

$$\frac{dT}{T}=\frac{2}{1-[(\gamma+1)/(\gamma-1)]Z}\left[\left(\frac{1-Z}{2}-\frac{\gamma Z}{\gamma-1}\right)d\theta-Zd\mu-Zd\alpha\right]. \quad (66)$$

Crystalline Aggregation of Cobalt Powder[a]

J. T. McCartney[b] and R. B. Anderson[c]

(Received April 28, 1947)

In electron microscopic studies of Fischer-Tropsch catalysts, an interesting phenomenon was observed in cobalt metal powder reduced from cobaltous oxide. The oxide particles sintered into larger smooth droplets of cobalt that were aggregated into thin hexagonal-shaped platelets. X-ray diffraction analysis showed the presence of the hexagonal crystal phase of cobalt. A possible explanation is that the aggregates are formed by forces similar to those operating in normal crystal formation, but of reduced magnitude insufficient to destroy the identity of the component particles.

IN the course of examination, with the electron microscope, of various components of cobalt catalysts used in the Fischer-Tropsch synthesis of liquid and solid hydrocarbons from mixtures of carbon monoxide and hydrogen, an interesting phenomenon was observed in cobalt metal powder reduced from the oxide. The cobalt was prepared by precipitation of cobaltous oxide from cobaltous nitrate with aqueous ammonia, reduction of the oxide in hydrogen at 250°C, and stabilization of the metallic cobalt against rapid oxidation by exposure to carbon dioxide at liquid nitrogen temperature. Figures 1 and 2 are electron micrographs showing the relative sizes and shapes of the particles of the oxide and the reduced metal. Surface area measurements by a nitrogen adsorption method[1] gave values of 67 square meters per gram for the oxide and 2 square meters per gram for the reduced cobalt. Thus, during reduction, there was an evident sintering of the irregular particles of the oxide into larger smooth droplets of cobalt metal.

Such droplets have been observed in several other metals prepared and reduced in a similar way. In cobalt, however, these particles are more or less closely aggregated into hexagonal platelets such as shown in Figs. 3 and 4. Varying degrees of dispersion in different specimens are illustrated in Figs. 2–4. The hexagons vary in diameter from 7 to 25 microns, and their thickness is apparently of the order of the diameter of the component particles. The hexagonal shape can be discerned also in the optical microscope. X-ray diffraction analysis shows that the crystal phase present is hexagonal cobalt. The hexagons are quite stable,

[a] Contribution from the Central Experiment Station, Pittsburgh, Pennsylvania. Published by permission of the Director, Bureau of Mines, U. S. Department of the Interior.

[b] J. T. McCartney, Physicist, Coal Constitution and Miscellaneous Analysis Section, Bureau of Mines, Pittsburgh, Pennsylvania.

[c] R. B. Anderson, Physical Chemist, Office of Synthetic Liquid Fuels, Research and Development Division, Bureau of Mines, Pittsburgh, Pennsylvania.

[1] S. Brunauer, P. H. Emmett, and E. Teller, J. Am. Chem. Soc. 60, 309 (1938).

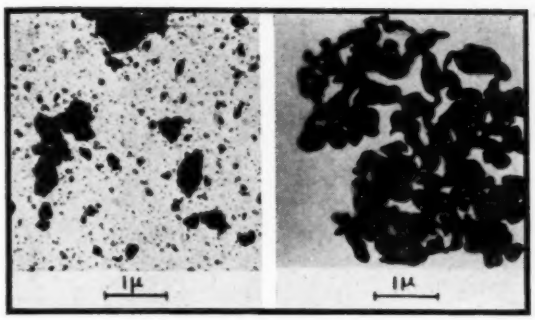


FIG. 1. FIG. 2.

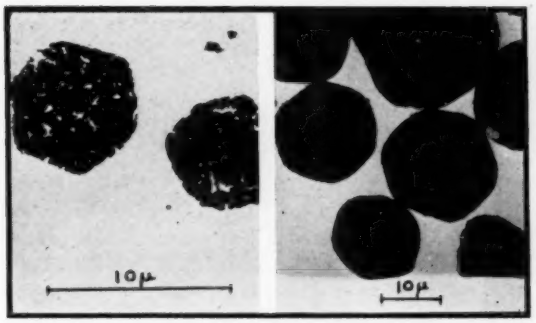


FIG. 3. FIG. 4.

FIGS. 1–4. Electron micrographs of cobaltous oxide (Fig. 1) and of cobalt metal reduced from it, in various stages of dispersion (Figs. 2–4).

their presence being observed in reduced cobalt samples which were at least partly oxidized and in reduced cobalt partially converted to cobalt carbide.

A possibly related phenomenon is that described by Picard and Duffendack[2] in evaporated metals. Their electron micrographs of evaporated zinc and cadmium show hexagonal crystals of the order of 0.1 micron in diameter. These are explained as resulting from migration and combination of the condensed atoms of the metals. In the Bureau of Mines laboratory, cobalt was evaporated onto collodion substrates and observed in the electron microscope. It apparently formed nearly continuous films with no observable migration into large crystals. Such was the case even at elevated temperatures, the cobalt condensing onto evaporated silica films heated probably above 250°C. Since cobalt has a much higher melting point than zinc and cadmium, higher temperatures may be required for migration.

It is suggested that the hexagonal aggregates in cobalt are formed by forces similar to those operating in normal crystal formation. Since, however, the component particles do not lose their identity and the aggregates are easily fractured, they can hardly be classed as true crystals. Possibly the forces acting to combine atoms into crystals are so reduced when the combining units are particles of relatively large size that only a superficial crystal formation occurs. It is not evident why this phenomenon has been observed only in cobalt.

[2] R. G. Picard and O. S. Duffendack, J. App. Phys. **14**, 291 (1943).

Surface Studies of Glass.

Part I. Contact Angles

L. A. SPITZE* AND D. O. RICHARDS
Physics Laboratory, Owens-Corning Fiberglas Corporation, Newark, Ohio
(Received May 5, 1947)

The water-repellent nature of glass surfaces treated with different silicones is examined by means of contact-angle measurements. Results are given showing the effect of various concentrations of silicones, types of glass, temperature of heat treatment, length of time of heat treatment, and surface roughness on the water contact angles. Silicones containing the methyl and phenyl group show good heat stability at temperatures of 500°C for short periods of time.

INTRODUCTION

IN applying glass to many industrial uses, more information is desired concerning its surface properties. This information is increasingly important as the surface area-weight ratio becomes progressively larger. In Fig. 1, the relationship between surface area and diameter of fibers is shown in English and metric units. From this graph it can be seen that fibers of glass are now being made having a surface area more than 50,000 times that of an equal unit weight of bulk glass. These extended surfaces present an opportunity for varied physical and chemical reactions to occur, and emphasize the need for an increasing amount of work on surface studies of glass. Such surface phenomena may possibly alter or mask many of the basic properties of the glass itself with the resultant properties being largely influenced by the type of surface presented. To learn more about some of these properties, different physical characteristics of the surface can be studied. A property of prime consideration is the degree of wetting of glass surfaces by liquids. Investigations of the degree of wetting of glass surfaces by liquids can be done readily by measurements of the contact angles these liquids form on the glass surface. Harkins and others[1] have shown that the contact angle for water, benzene, carbon tetrachloride, chloroform, and ether in contact with clean glass is zero. A solid is completely wetted by a liquid if the contact angle is zero, and incompletely wetted if a measurable contact angle is found. A zero contact angle indicates that the liquid attracts the solid as much as it attracts itself, a 90° contact angle shows that the attraction of the liquid for the solid is half that for itself, and a contact angle of 180° would mean there was absolutely no attraction of the liquid for the solid. Because contact angles of 180° are unrealizable, complete non-wetting is impossible; but solids with liquid contact angles greater than 90° are generally referred to as being non-wetted. In such industrial applications of glass fibers as flotation mats, water-repellent fabrics, and glass-plastic laminates with decreased water sorption, a hydrophobic surface is desired. Such materials as the long-chain fatty acids or their salts, when applied to a glass surface, will make it hydrophobic. Although these materials will give a hydrophobic surface, they possess poor durability under various conditions, and particularly poor stability at elevated temperatures. Recently a new group of materials has been introduced containing members that also have the ability to give a hydrophobic surface to glass when treated by them. These materials, commonly referred to as the silicones,[2] have the distinct advantage of good heat stability. Materials such as the silcones, when applied to glass surfaces, permit the use of much higher temperatures in the applied product than would be available for glass surfaces treated with the usual organic materials. Thus, more of the inherent heat-resistant properties of the glass will be available for use in products containing glass

* Present address: Department of Chemistry, Southwestern College, Winfield Kansas.

[1] Harkins and Brown, J. Am. Chem. Soc. **41**, 499 (1919); Richards and Carver, J. Am. Chem. Soc. **46**, 1196 (1924).

[2] C. A. Scarlott, Westinghouse Engineer **5**, 130 (1945); H. Hausman, J. Chem. Ed. **23**, 16 (1946); H. A. Knight, Materials and Methods **23**, 1069 (1945); S. L. Bass, Elec. Eng. **66**, 381 (1947).

treated with these materials. In addition, these silicones, being largely inorganic materials containing silicon, could be expected to be compatible with a clean glass surface and would firmly adhere to such surfaces.

The degree of non-wetting that the silicones impart to a glass surface depends on the individual member, and on the method by which the silicone is applied to the glass surface.

EXPERIMENTAL

The hydrophobic nature of a glass surface treated with various silicones was studied by determining the water contact angles. Measurements of contact angles may be made by various procedures. Methods most commonly used include direct measurement of the angle by use of photography or projection, tilting plate,[3] sessile drop,[4] or pressure of displacement in capillary tubes. The method selected for this study was the sessile drop. Determinations are made by placing a number of small drops of water on the treated glass surface and then measuring the dimensions of the drops with the aid of a microscope. The size of the contact angle, θ, is then calculated from the relationship[5] $\tan\theta/2 = 2h/d$ where h is the height of the drop, and d is the width of the drop at the place of contact with the solid. The sessile drop method can be used only if the portion of the drop measured is a segment of a sphere. It has been shown that if the size of the drops is kept smaller than 1 mm in diameter, the contact angle attains a constant maximum value that is the advancing contact angle.[4] This method was selected for this study because of the ease of adapting the available equipment, and measurements of a number of drops on a treated surface will give an average value for the drops placed at various positions on the treated surface.

The apparatus used to determine the contact angles of water on the treated glass surfaces consisted of a Bausch and Lomb microscope placed in a horizontal position and containing a 10× objective. The drop was viewed and measured by means of a Bausch and Lomb 12.5× Filar eyepiece. Most of the glass samples used were cut to about 1″×½″ by ¼″ thick. The smooth, uncut, fire-polished upper surface was used in the studies. The treated glass sample was placed on a stage and brought into the line of sight by a mechanical microscope stage. A Spencer microscope illuminator was used to give a sharp outline of the drops placed on the glass surface. Drops of distilled water of fairly constant size, less than 1 mm in diameter, were placed on the treated glass surfaces by means of capillary tubes. For some of the treated surfaces having high contact angles, it became necessary to place the drops on the surface by spraying, as the adhesion tension of the water to the glass capillary was much greater than that for the treated glass surface. No fewer than five separate drops in the desired size range were measured and contact angles calculated. The glass samples were first cleaned with chromic acid cleaning

TABLE I. Contact angles of water on trichloro-organosilane treated *E*-glass.

15 Min. @ °C	30°	75°	125°	200°	300°	400°	500°
Trichloromethylsilane	97	99	104	103	101	96	94
Trichlorophenylsilane	81	93	100	101	97	91	84
Trichloroethylsilane	92	100	103	97	90	60	0
Butyltrichlorosilane	96	98	101	96	90	60	—
Trichlorooctylsilane	98	99	101	96	86	0	—
Trichlorododecylsilane	111	108	107	103	84	35	0
Trichlorooctadecylsilane	106	106	103	70	66	10	0

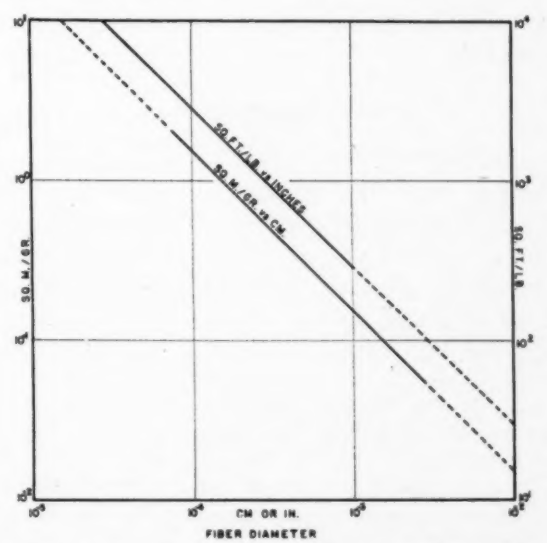


FIG. 1. Surface area *vs.* fiber diameter; solid lines show present working range in the fibrous-glass industry.

[3] N. K. Adam and G. J. Jessop, J. Chem. Soc. (London) **127**, 1863 (1925); E. L. Green, J. Phys. Chem. **33**, 921 (1929); R. N. Wenzel, Ind. Eng. Chem. **28**, 988 (1936).
[4] G. L. Mack, J. Phys. Chem. **40**, 159 (1935).
[5] F. E. Bartell, *Laboratory Manual of Colloid and Surface Chemistry* (Edwards Brothers, Inc., Ann Arbor, Michigan, 1938), p. 85.

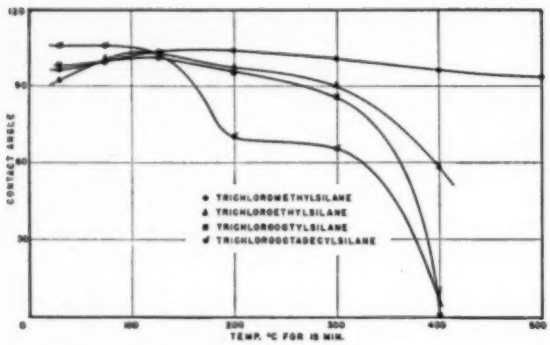


FIG. 2a. Contact angles of water on *E*-glass treated with 1 percent solution of trichloroorganosilanes *vs.* temperature.

solution and then by flaming. The cleaned glass was then dipped in a 1 percent solution of the silicone in c.p. benzene, removed and allowed to drain a short while, and then placed in an oven at the desired temperature. After the sample was in the oven the desired time, it was removed and allowed to cool. The treated samples were then placed on the microscope stage and the dimensions of the drops placed on them were measured by use of the Filar eyepiece in the microscope.

RESULTS

Values of the contact angles of water were obtained for glass surfaces treated with organosilane chlorides and with organosiloxane polymers. The magnitude of the contact angle was found as a function of the variables of temperature, length of time of heat-treatment, type of glass, and concentration of the surface treatment.

Trichloroorganosilanes

Various members of this particular group of silicones were used to treat the glass surfaces. These materials were diluted with c.p. benzene until a solution was obtained containing 1 percent solids. Contact angles were obtained for these materials, first as a function of the temperature, the other variables being held constant. The samples of glass used here were a borosilicate glass type containing almost no alkali (Na_2O or K_2O) and designated as *E*-glass by Owens-Corning Fiberglas Corporation. After being treated with the silicone solutions, the samples were placed in the oven at the desired temperature and left for fifteen minutes. The results of these measurements are shown in Table I. The data of Table I have also been presented graphically in Figs. 2a and 2b. These values represent determinations made at different times on different samples of the same type of glass. At temperatures below 300°C the results were most consistent. At the higher temperatures, the results varied somewhat. This can be expected from the shape of the curves as shown in Figs. 2a and 2b. Because the curves are steep in this region, any small variation in heating temperature or length of time of heating would greatly affect the results.

TABLE II. Contact angles of water on *E*-glass treated with organosiloxane polymers.

15 Min. @ °C	30°	75°	125°	200°	300°	400°	500°	550°
Dimethylsiloxane polymer	31	45	67	92	101	99	96	31
Diethylsiloxane polymer	40	46	82	102	97	91	44	25
Methylphenylsiloxane 0.5 Ph/Si	31	43	49	90	101	97	97	80
Methylphenylsiloxane 0.9 Ph/Si	25	44	61	91	91	90	91	91

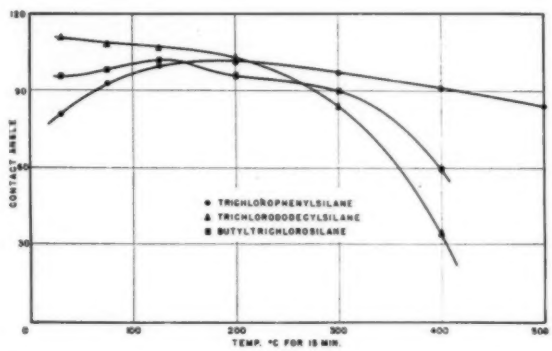


FIG. 2b. Contact angles of water on *E*-glass treated with 1 percent solution of trichloroorganosilanes *vs.* temperature.

The effect of the type of glass being treated on the contact angle is shown in Fig. 3. Three different types of glasses were used, treated with trichlorooctadecylsilane (1 percent solids), and heated at the specified temperatures for fifteen minutes. The three glass types used were *E*-glass, Pyrex chemical glass, and window glass. Chemical analysis shows that these glasses contain different amounts of alkali (Na_2O and K_2O). The *E*-glass has less than 1 percent, the Pyrex chemical glass about $4\frac{1}{2}$ percent, and the window glass about $12\frac{1}{2}$ percent of these alkali materials.

Organosiloxane Polymers

In Table II are shown the results of contact-angle measurements of water on *E*-glass treated with 1 percent solutions in benzene of four different organosiloxane polymers. These materials were heated at the desired temperature for fifteen minutes. The data have been plotted in Fig. 4 for visual use of these results.

It was considered advisable to determine the effect of different periods of heating on the contact angles of glass surfaces treated with these polymers. In Table III is shown the effect of various heating periods up to six hours on *E*-glass. Three different temperatures were used for this study: room temperature, 300°C, and 450°C. An upper temperature of 500°C was tried at the beginning, but for periods greater than one-half hour the results varied widely. Consequently, this "burning off" region was avoided and a lower temperature of 450°C selected. The results obtained are also shown graphically in Figs. 5–7.

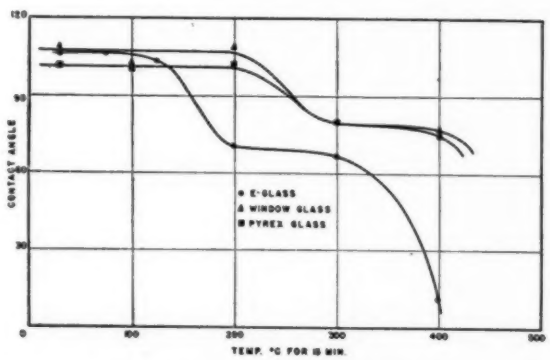


FIG. 3. Contact angles of water on various glasses treated with 1 percent solution of trichlorooctadecylsilane *vs.* temperature.

In addition, the effect of various concentrations of the silicones on *E*-glass was studied by contact-angle measurements. A curve showing the effect of dimethylsiloxane polymer on *E*-glass heated at 200°C and at 500°C for 15 minutes is shown in Fig. 8. A trichloroorganosilane was also used to study its effect at different concentrations. The results obtained for this material are discussed later.

DISCUSSION

Before making contact-angle measurements by the sessile drop method on silicone-treated glass,

TABLE III. Contact angles of water on *E*-glass treated with organosiloxane polymers.

Material	Temp. °C	Heating time in minutes 15	30	60	180	360
Dimethylsiloxane polymer	30	31	37	38	40	46
	300	101	104	104	100	—
	450	101	93	93	81	81
Diethylsiloxane polymer	30	40	53	54	55	55
	300	97	101	93	87	—
	450	89	79	62	62	20
Methylphenylsiloxane 0.5 Ph/Si	30	38	34	35	36	34
	300	101	96	99	96	—
	450	93	99	90	82	56
Methylphenylsiloxane 0.9 Ph/Si	30	25	45	39	39	36
	300	91	90	96	93	—
	450	96	99	93	79	86

the method was tested by comparing the results obtained on some common materials with the values of the contact angles for these materials as reported in the literature. Measurements made of the contact angle of water on paraffin gave an average value of 105°, which compares favorably with the values reported elsewhere.[6] A photo-micrograph of a drop of water on paraffin is shown in Fig. 9. From this photograph it can be noted that the drop is a segment of a sphere, and the sessile drop method of measurements will give the contact angle for drops this size.

Two additional photo-micrographs of water on *E*-glass treated with dimethylsiloxane polymer are shown in Figs. 10 and 11. Figure 10 shows the appearance of water on *E*-glass treated with the polymer and heated to 125°C for 15 minutes. After the treated glass has been heated to 300°C for 15 minutes, the appearance is shown in Fig.

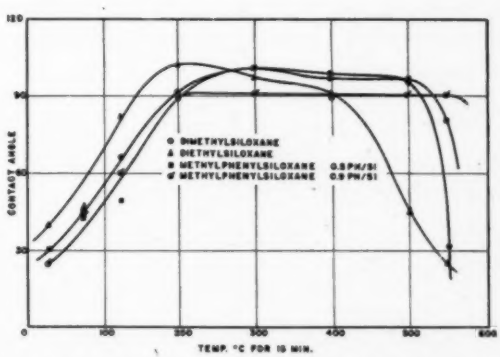


FIG. 4. Contact angles of water on *E*-glass treated with 1 percent solution of organosiloxane poymers *vs.* temperature.

[6] Int. Crit. Tab. **4**, p. 434 (1928); N. K. Adams, *The Physics and Chemistry of Surfaces* (Oxford University Press, London, 1941), p. 186.

11. This increase of heat raises the contact angle 34°.

Observation of the data in Table I and the curves of Figs. 2a and 2b bring out the superior heat stability of the trichloroorganosilanes containing the methyl and the phenyl groups. These data show that the heat stability as determined by water repellency of these trichloroorganosilanes is dependent upon the particular organic group attached to the silicon. From these results it appears that the heat stability of the organic groups tested may be arranged in this descending order: methyl, phenyl, ethyl, butyl, octyl, dodecyl, and octadecyl. The maximum water repellency is shown by the materials containing the dodecyl and the octadecyl groups. Although the data in Table I do not show a definite order, it is most likely that the arrangement of the organic groups in the order of decreasing water repellency would be the reverse of the order for heat stability as given above. In particular, the methyl group seems to deviate from this arrangement. The maximum contact angles shown by these materials are all within a 10° range; hence any experimental errors may account for the fact that the data are not in the orderly arrangement suggested above. The data for the organosiloxane polymers again show the heat stability of the methyl and the phenyl group. For these materials, the polymer containing the highest ratio of phenyl to silicon shows the best heat stability. Although there is very little difference in the values for the maximum contact angles obtained, the polymer containing the ethyl group appears to have a slightly higher maximum contact angle.

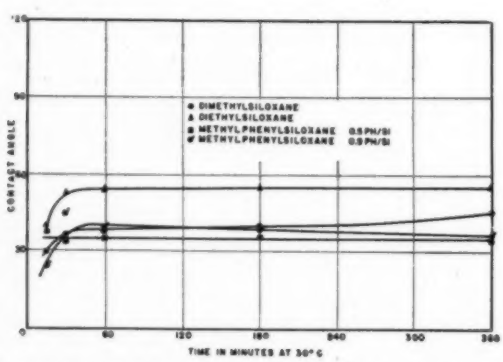


FIG. 5. Contact angles of water on *E*-glass treated with 1 percent solution of organosiloxane polymers *vs.* heating time.

The curves of Fig. 3 show a noticeable difference in the contact angles of various types of glasses treated with trichlorooctadecylsilane. These curves show that the highest contact angles are obtained for the glass containing the most alkali. Thus it is believed that the higher contact angles for window glass and for Pyrex glasses are due to the higher alkali content. It appears that the silicone materials will attach themselves more firmly to the alkali glasses, possibly because of a higher surface activity of these glasses. Results obtained when a dimethylsiloxane polymer was used to treat these various glasses did not show such a difference as exhibited by the trichlorooctadecylsilane-treated glasses. The two treated glasses containing the most alkali did, however, show appreciably higher contact angles at treatment temperatures below 200°C and values a little higher at temperatures above 500°C. It is possible that the polymer is able to orientate itself somewhat faster on these glasses containing alkali and that it is also bound a little stronger to these glasses.

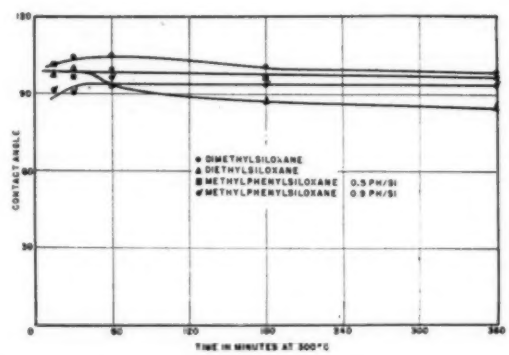


FIG. 6. Contact angles of water on *E*-glass treated with 1 percent solution of organosiloxane polymers *vs.* heating time.

The mechanical condition of the surface of the treated glasses influences the wettability to a large extent. The effect of the roughness of a surface on the contact angles has been recognized[7] as due to a greater net energy change in the wetting process caused by a rough surface. A drop tending to spread on a rough surface will increase the free liquid surface, but the wetted area under the drop will increase to a larger extent because of the rough surface. Increase of

[7] R. N. Wenzel, Ind. Eng. Chem. **28**, 988 (1936).

the liquid surface involves a consumption of energy, while the increase of wetted area under the drop involves a release of energy. The difference is a net energy change which determines the speed of wetting on the solid. For a rough solid tending to be wet by a liquid, a greater net energy decrease is present than if the surface were smooth and this causes the rough surface to be wetted more thoroughly. For a liquid-repellent rough surface, the same reasoning can be applied. For a change in the shape of a drop on a rough liquid-repellent surface, more actual surface change is involved at the liquid-solid interface. As the dry interface has the lower specific energy, the net energy decrease is greater

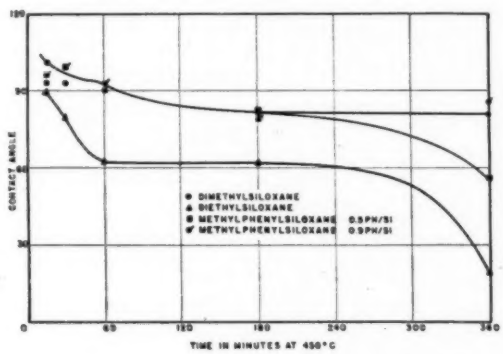


FIG. 7. Contact angles of water on *E*-glass treated with 1 percent solution of organosiloxane polymers *vs.* heating time.

for the rough surface causing the drop to assume a more spherical form. A roughened surface magnifies the wetting properties of a solid, making it either more liquid-repellent or more easily wetted.

The effect of a roughened glass surface on the contact angles was investigated briefly. A sample of *E*-glass was roughened by grinding with an abrasive and another sample was roughened with an acid etch. These surfaces were cleaned and then treated with a 1 percent solution of dimethylsiloxane polymer and heated for 15 minutes at 300°C. The contact angles obtained on these samples gave values of from 120° to 125°, depending somewhat on the manner by which the surface was roughened. A photo-micrograph of the appearance of a drop of water on a roughened *E*-glass surface treated with dimethylsiloxane polymer at 300°C for 15 minutes

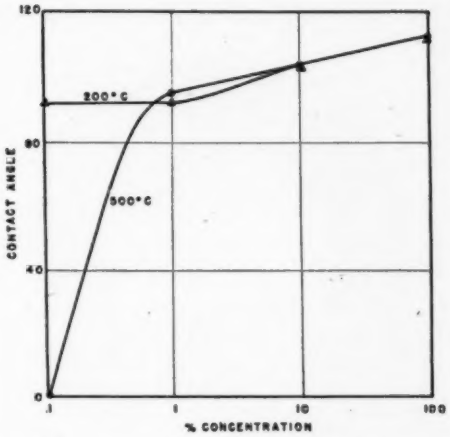


FIG. 8. Contact angles of water on *E*-glass treated with dimethylsiloxane polymer at 200°C and 500°C *vs.* concentration.

is shown in Fig. 12. The contact angle for this sample was found to be 124°, or an increase of 23° over the value of a treated smooth surface. The industrial applications of the surface roughness of materials, as an aid to water-repellency, are being studied and evaluated at the present time.

The effect of various lengths of time of heating these treated glass surfaces is small at room temperature. The data in Table III and Figs. 5–7 show, also, that at 300°C there is a small effect, if any, because of variation in length of time of heating these treated samples. Two of the samples appear to have a slightly larger contact angle as the time is increased to 30 minutes, but the effect is small. At 450°C, a more noticeable effect begins to appear. At this higher temperature, some of the material is beginning to burn off. This is particularly so for the diethylsiloxane polymer, and to a lesser extent for the methylphenylsiloxane polymer (0.5 Ph/Si). The data for the diethylsiloxane polymer also show a drop in contact angle with time up to 60 minutes, then remain steady until it begins to burn off. This drop may be due to the fact that equilibrium for this material is not reached at the higher temperatures until after 60 minutes. This effect is also somewhat noticeable for the dimethylsiloxane polymer.

The effect of different concentrations of the silicones on glass is significant, particularly at the lower concentrations. The data in Fig. 8 show the effects of varying the concentration of

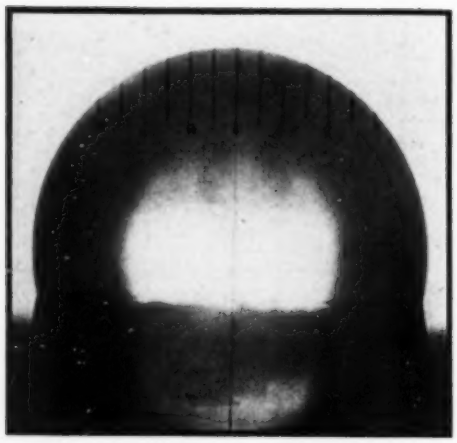


FIG. 9. Appearance of a drop of water on a paraffin surface; contact angle 105°.

dimethylsiloxane on *E*-glass and then heating at 200° and at 500°C for 15 minutes. The data show that an increase in concentration raises the contact angle at a particular temperature. This is most pronounced at the higher temperature and the lower concentration. At 500°C the 0.1 percent solution burns off, but the 1 percent solution gives a high contact angle. The 0.1 percent solution, although heated at 200°C, is not destroyed and gives a high angle. Although experiments were not tried at lower concentrations, it is believed that lower concentrations of dimethylsiloxane polymer would yield lower contact angles for the 200°C curve. Thus if these experiments had been conducted at lower concentrations, the resultant contact angle *vs.* temperature curves would very probably have been lower and burned off at considerably lower temperatures. The increase of the contact angles resulting from increase of concentration may be due in part, at the higher temperatures, to an increased surface roughness. This roughness is probably caused by deposition of some SiO_2 on the glass by the high heat, and any undestroyed silicone will give the necessary water-repellency necessary to obtain the higher contact angles. The concentration effect in the case of the trichloroorganosilanes is not as pronounced as that of the organosiloxane polymers. However, at the higher temperatures and lower concentrations, a noticeable drop in the values for the contact angles was apparent. These materials would also be expected to give values of contact angles somewhat lower as the concentration is decreased, but not to such an extent as that shown by the organosiloxane polymers.

Water-repellent glass in the form of fine fibers has proved to be very useful. For example, these treated glass fibers have proved their superiority over other materials as buoyant materials for Navy life preservers.[8] This superiority is due to their fire- and mildew-proofness, superior buoyancy, ability to withstand repeated wetting and drying with little loss in efficiency, and the fact that the treated glass fibers can be produced in

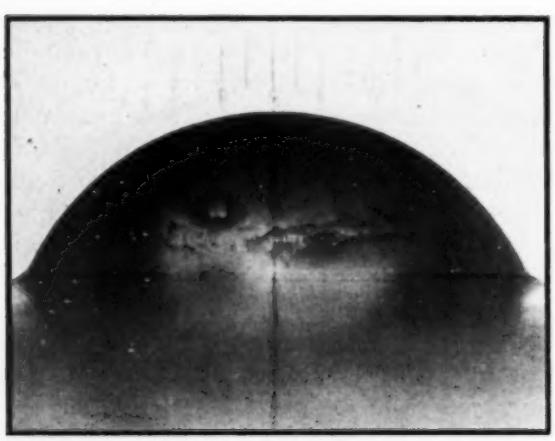


FIG. 10. Appearance of a drop of water on an *E*-glass surface treated with a 1 percent solution of dimethylsiloxane polymer and heated for 15 minutes at 125°C; contact angle 67°.

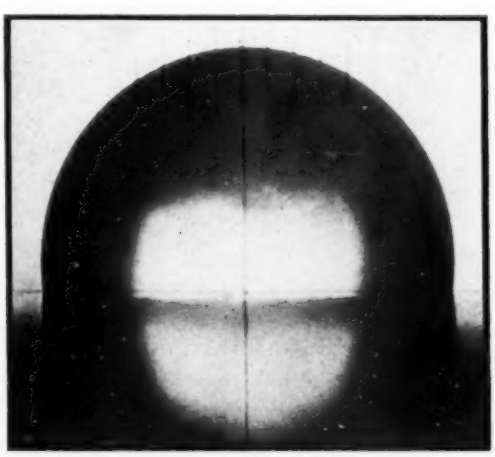


FIG. 11. Appearance of a drop of water on an *E*-glass surface treated with a 1 percent solution of dimethylsiloxane polymer and heated for 15 minutes at 300°C; contact angle 101°.

[8] A. L. Holden, U.S.N. Inst. Proc. **72**, 1327 (1946).

uniform quality. Such treated glass fibers have been adopted by the Navy as standard material.[8] In addition, these silicone-treated glass fibers have shown their value in reinforced plastics.[9] Their water-repellent properties have been very useful in increasing wet strength of these laminates. The wet strength of cloth laminates, improved by the application of a silicone solution to the glass fibers, has been brought up to a point nearly equal to the dry strength.[10]

This paper shows the results of contact-angle measurements of only a few of the many possible types of silicone materials. More information obtained on the other types of silicone materials available, when applied to glass and other materials, will be of utmost importance. Studies other than contact-angle measurements will also prove their worth in evaluation of surface properties of surfaces treated by the silicones.

CONCLUSIONS

The values of the contact angles of water on glass surfaces treated with various trichloroorganosilanes and organosiloxane polymers show the good water-repellency properties these materials bestow on a glass surface. In addition, the superior heat stability of these materials is demonstrated. Of the various organic groups in the silicone molecule, the methyl and phenyl impart the best heat-resistant properties, and the dodecyl and octadecyl give the materials their best water-repellency. The value of increased surface roughness in giving better water-repellency also is shown. Finally, the effect of variations in the concentrations of these silicones when applied to the surfaces is given. It is recommended that concentrations no lower than 1 percent solids be used when maximum heat stability and water-repellency is desired.

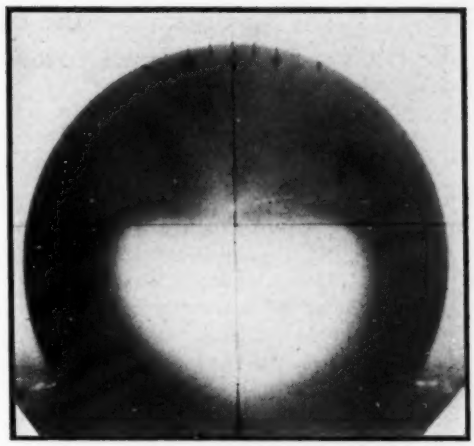


FIG. 12. Appearance of a drop of water on a roughened surface of *E*-glass treated with 1 percent solution of dimethylsiloxane polymer and heated for 15 minutes at 300°C; contact angle 124°.

ACKNOWLEDGMENT

The silicone materials used in this study were furnished by Dr. M. J. Hunter of Dow-Corning Corporation, Midland, Michigan. The authors also wish to acknowledge the aid given during the course of this study by R. B. Taylor, W. C. Marlow, L. P. Biefeld, Robert Steinman, and G. F. Forman, all of the Research Division of Owens-Corning Fiberglas Corporation.

[9] P. L. Alger, Gen. Elec. Rev. **50**, 12 (1947).
[10] G. F. Nordenholt, Prod. Eng. **18**, 160 (1947).

An Apparatus for the Observation of the Kerr Effect with Microsecond Electrical Pulses

WILBUR KAYE AND RICHARD DEVANEY
Tennessee Eastman Corporation, Research Laboratories, Kingsport, Tennessee
(Received May 9, 1947)

An apparatus for the observation of the Kerr electro-optic effect in liquids and polymers using microsecond electrical pulses, is presented and evaluated. This method offers a convenient way of studying internal motions and molecular configurations of polymeric material with a minimum of heating usually accompanying these observations. Electrical pulses of from 1- to 1000-microsecond width, 60–1000 p.p.s., at amplitudes up to 5000 volts are generated by conventional methods. These pulses are used to orient low molecular weight and polymeric materials placed between the plates of a condenser. The circular polarization induced in the polarized light passing between the plates actuates a photo-multiplier tube in order that the light pulses may be studied with an oscillograph.

Curves of Kerr constant *versus* temperature at temperatures to −80°C are compared with the literature. Oscillograms showing the response of several liquids and polymers are shown. Oscillograms showing an interesting difference between orientation and relaxation time are presented, as are oscillograms showing clearly both orientation and distortion polarization simultaneously.

BRIEFLY the Kerr electro-optic effect is concerned with the elliptical polarization induced in a material subjected to an electrical field.[1] It is the aim of this paper to present the instrumental details and the evaluations necessary to utilize this Kerr effect for the study of internal motions and molecular configurations of organic liquids and polymers.

The instrument described in this report utilizes pulsed electrical fields for the study of the Kerr effect. This method offers the advantage of minimizing heating effects in the material under investigation, and permits one to study the nature of the orienting effect in a more direct manner than has hitherto been accomplished by the use of sinusoidal electrical fields.[2] A circuit similar to that used here was described by Pauthenier[3] using pulsed condenser discharges. However, Pauthenier had no direct method of studying the shape of the response wave as is now practical by use of the oscillograph.

Figure 1 shows a photograph of the exterior of the apparatus, which in conjunction with Fig. 2, a block diagram, will be used to explain the methods by which the Kerr effect may be produced and observed.

Electrical pulses of the order of a few microseconds width are generated every hundredth of a second by the pulse generator. The high voltage amplifier then boosts the voltage of the pulses to several thousand volts amplitude without serious distortion. This electrical pulse on the plates of the Kerr cell causes the plane-polarized light passing through the sample to be circularly polarized; hence, it passes through the analyzing Nicol which is crossed with respect to the polarizing Nicol and oriented 45° with respect to the field in the Kerr cell. The light pulse subsequently produced by the electrical field actuates the photo-tube, wherein it is converted into an electrical signal that can be amplified and studied with the oscillograph.

Ordinary low molecular weight compounds are capable of orienting in an electrical field within 10^{-9} second;[4] hence, the wave shape of the optical pulse is very nearly the same as that of the electrical pulse, provided certain restrictions of wave form and amplitude are made. On the other hand, there are materials composed of large molecules in a viscous medium which do not follow the electrical field at such a rapid rate. The optical pulse from such material should be expected to show considerable distortion in relation to the electrical field if one assumes a molecular orientation effect to be responsible for the electro-optic effect.

[1] J. Kerr, Phil. Mag. **50**, 337, 446 (1875).
[2] J. W. Beams, Rev. Mod. Phys. **4**, 133 (1932).
[3] M. J. Pauthenier, J. de phys. et rad. **2**, 183, 384 (1921).
[4] J. W. Beams and Lawrence, J. Frank. Inst. **206**, 109 (1928).

The amount of light passing the analyzing Nicol is related to the electrical field intensity, F, and the optical path length, l, between the Kerr cell plates by the expression[5]

$$I=a^2\sin^2(\pi KlF^2)=a^2\sin^2\delta/2, \qquad (1)$$

where K is the Kerr constant and δ the angular retardation. The general relationship of K with other molecular factors is[6]

$$K=\frac{(n^2-1)(n^2+2)}{4n^2}\,\frac{(\epsilon+2)^2}{(3)}\,\frac{\theta_1+\theta_2}{\alpha}, \qquad (2)$$

where n is the refractive index, ϵ, the dielectric constant, α, the polarizability, and θ_1 and θ_2 are defined as

$$\theta_1=(1/45kT)[(a_1-a_2)(b_1-b_2)+(a_2-a_3)(b_2-b_3)+(a_3-a_1)(b_3-b_1)],$$
$$\theta_2=(1/45k^2T^2)[(\mu_1^2-\mu_2^2)(b_1-b_2)+(\mu_2^2-\mu_3^2)(b_2-b_3)+(\mu_3^2-\mu_1^2)(b_3-b_1)].$$

Here, k is Boltzmann's constant; T, the absolute temperature; a_1, a_2, and a_3 are proportional to the shape of the induced polarization ellipsoid of the molecule in three mutually perpendicular directions; b_1, b_2, and b_3 are similar coefficients of the optical ellipsoid of the molecule; while μ_1, μ_2, and μ_3 are the components of the permanent dipole moment. The importance of the temperature dependance arising from thermal motion will be discussed later.

From what has been said above, it should appear possible to find a material which would show a definite lag at easily obtainable temperatures in response to the electrical field, and that lag would be markedly dependent on temperature. Such a material would have to be composed of large polar molecules in a viscous medium. These requirements are met by a number of synthetic plasticizers and polymers. The literature records the observation of such lags in rosin,[7] glass,[8] benzopurpurin,[9] and vanadium pentoxide solutions.[10]

[5] E. F. Kingsbury, Rev. Sci. Inst. 1, 27 (1930).
[6] P. Debye, Handbuch Radiologie 6, 754 (1925).
[7] J. Kerr, see reference 1; D. W. Kitchin and H. Muller, Phys. Rev. 32, 979 (1928).
[8] J. Kerr, see reference 1.
[9] J. Errera, J. Overbeek, and H. Sack, J. de Chimie Phys. 32, 681 (1935).
[10] J. Errera, J. Overbeek, and H. Sack, see reference 9.

GENERATION OF PULSE

In this section use will be made of Fig. 3, a schematic diagram of the pulse amplifier, to give a brief description of its operation. The pulse generator is a Measurements Corporation model 79-B unit. The output circuit of this generator has been modified somewhat to give control of amplitude and output impedance as well as modified to obtain wide pulses. Normally, the generator is used with full amplitude (approximately 150 v) and 1000-ohm output impedance. Pulses from $\frac{1}{2}$ to 1000 microseconds and repetition rates from 60 to 100,000 p.p.s. are obtainable.

The output coupling condenser has been removed for certain applications to be mentioned later. However, the coupling condenser is in the circuit located on the amplifier chassis to permit control of pulse amplitude through the grid bias. Such a method of amplitude control was used to permit the design of twin amplifiers using a common signal source and plate supply. The amplifier tubes are Eimac 4–125A tubes. Two pairs of load resistors are available, a 10,500-ohm load for narrow pulses and a 75,000-ohm load for wide pulses.

Taps in the load resistors permit a portion of the output signal to be deflected directly to the plates of the cathode-ray tube for observation of

FIG. 1. Exterior of apparatus. Pulse generator in left rack, optical bench, amplifier, and oscillograph on bench and dielectric bridge in right rack.

the electrical wave form. This circuit attenuates the output signal but not the a.c. ripple; hence, it greatly enhances this ripple.

Another tap in one of the load resistors permits the use on the Kerr cell of an attenuated pulse, improving the wave form when both channels are operated simultaneously, one with considerably greater amplitude than the other.

The peak pulse voltage is measured by means of a Sensitive Research Instrument Corporation model ESH multi-range electrostatic voltmeter. The pulses are isolated by means of a diode as shown. Care has to be taken in this circuit to place a high resistance to ground to prevent the accumulation of a d.c. charge formation through leakage of the coupling capacitor. The two amplifier channels are carefully shielded for accurate voltage determinations. Switching is accomplished by means of Eimac vacuum switches in the high voltage circuits and antenna relays in the low voltage circuits.

Figure 4 is an oscillogram of the electrical pulse of 10-microseconds width. The wave form has been improved by driving the circuit to saturation in the sense that the pulse voltage approximates the plate supply.

OPTICAL SYSTEM

The light source mounted on a Cenco optical bench is a Spencer 6-volt lamp operated from a battery to avoid any periodic fluctuations of light intensity. Cenco 10-mm Nicol prisms act as polarizing and analyzing media. The Kerr cells are mounted ridgidly in line. Figure 5 shows the construction of one of these cells and its temperature bath which has operated successfully to keep the temperature constant to within one degree from -80 to $+100$°C, using a dry ice ether mixture in the side chambers and an immersion heater with De Kotinsky thermo-regulator to maintain the temperature in an alcohol or oil medium. The light enters and leaves through double windows maintained clear at low temperatures by a stream of air. Considerable difficulty was encountered before suitable cell plates were found which could be kept parallel at separations less than a millimeter and still be small and easily cleaned after use with polymeric materials. A solution to this problem was found by coating ordinary glass microscope slides with a layer of gold deposited from duPont Liquid Bright Gold and fired at 1100°F for a short time. Both sides of the slides are coated, as are three of the edges, and electrical contact is made by spring contacts to the outside surface of the slides. The slides are separated by mica spacers averaging 0.016 in. Thicker glass spacers have been used for materials of high Kerr constant.

A Gaertner model L-135 Babinet-Soleil compensator is placed between the Nicols to permit measurement of the retardation, zero out small retardations in the glass windows, and increase the sensitivity of the system as the occasion may require. The compensator requires 4.42 turns per wave of light of 465-mu wave-length.

The analyzing Nicol is crossed with respect to the polarizing Nicol and at an angle of 45° with respect to the plane of the electrical field. A Baird interference filter passing light in the neighborhood of 456 mu was placed directly over the iris diaphragm acting as a window to the photo-tube chamber.

The type 931-A multiplier photo-tube was chosen because of its sensitivity and frequency response.[11] The voltage-divider system for the

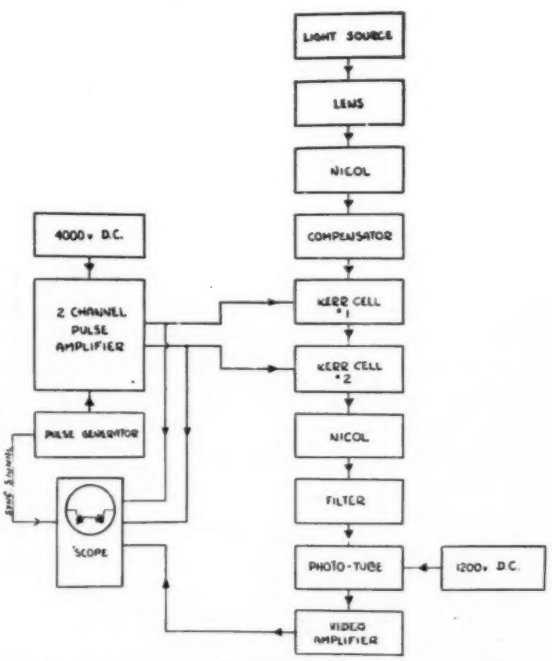


FIG. 2. Block diagram of Kerr effect apparatus.

[11] R. C. Winans and J. R. Pierce, Rev. Sci. Inst. **12**, 269 (1941); R. D. Rawcliffe, Rev. Sci. Inst. **13**, 413 (1942).

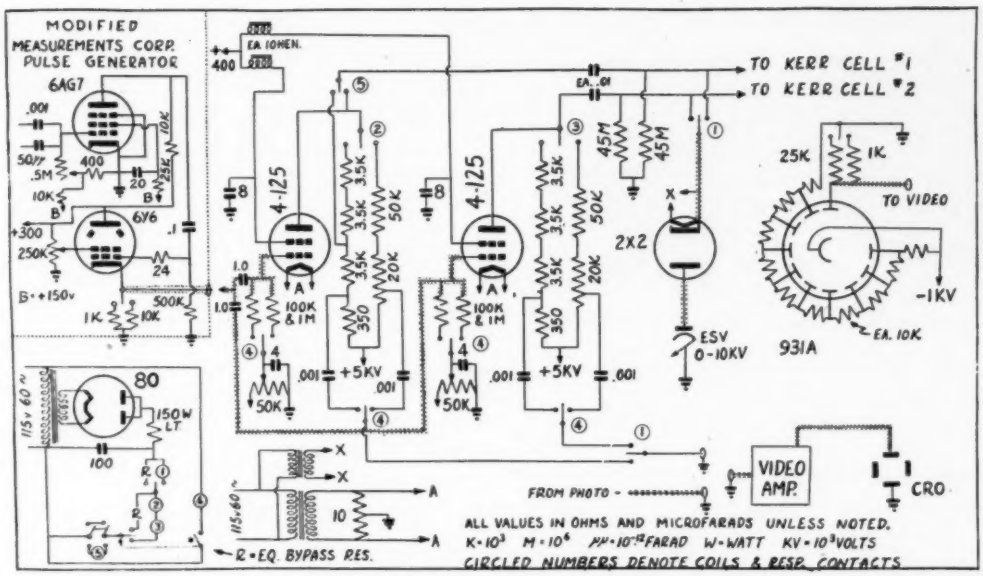


FIG. 3. Wiring diagram of Kerr effect apparatus.

multiplier dynodes is mounted in a shielded box directly below the photo-tube. The load resistor of the photo-tube is variable to permit control of the output amplitude and fidelity of the high frequency response. This resistor likewise is mounted close to the photo-tube and connected to the video amplifier by a short shielded lead.

The video amplifier is a United Cinephone model 8240 with a gain of 1000 and a frequency essentially flat from 15 cycles to 4 megacycles. The output of this video amplifier feeds directly to the plates of the cathode-ray tube in a DuMont 208B oscilloscope.

INSTRUMENTAL RESPONSE

An elaborate discussion of the optics of this type of system will not be given here since it may be found in the literature.[12] However, attention should be given to formula (1), applicable to the above conditions where it may be seen that the light intensity, hence deflection of the oscilloscope tracing, will be primarily a function of the field strength. With a material of sufficiently large Kerr constants it is possible to observe a sinusoidal fluctuation of light intensity on the leading and trailing edges of the pulse. Figure 6 shows an oscillogram of several superimposed optical pulses with nitrobenzene at low voltage. A comparison of this figure with that of the electrical pulse will furnish a measure of the performance of the instrument. Figure 7 is an oscillogram of the response of nitrobenzene to a large voltage pulse. For this photograph the electrical circuit was set to produce sloping edges and high amplitude. With reference to Eq. (1) it is seen that seven waves of retardation are produced. In order to obtain such a pattern, it is necessary to use monochromatic light and an amplifier of high fidelity in addition to accurately parallel cell plates properly aligned.

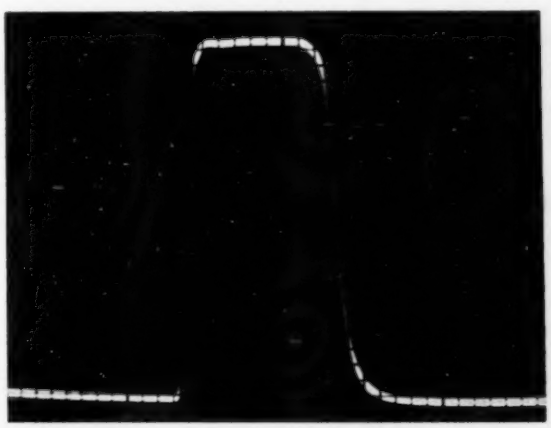


FIG. 4. Oscillogram of electrical pulse; 10-μsec. width, 650 p.p.s. 4500-volts amplitude.

[12] E. F. Kingsbury, see reference 5.

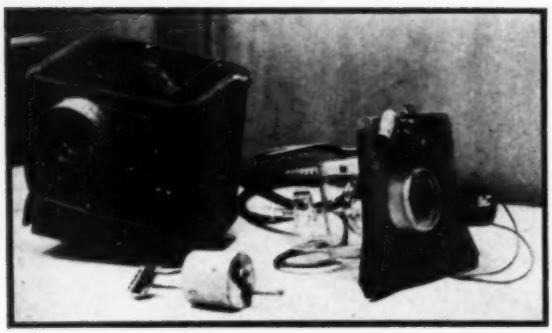


FIG. 5. Kerr cell and temperature bath.

METHODS OF MEASUREMENT

The simplest method of measuring the Kerr constant would be to measure the amplitude of the signal and compare this amplitude with a material of known Kerr constant. This method, however, is inaccurate though useful for the estimation on materials of very low Kerr constant.

A method of greater accuracy uses the Babinet-Soleil compensator to produce a retardation equal and opposite to that produced in the Kerr cell. The procedure is as follows.

The amplitude of the pulsed retardations produced by the Kerr cell is set to an arbitrary value by varying the gain of the video amplifier. The compensator is introduced and set to a value that will just *invert* the signal an equal and opposite amount. The difference in compensator readings from its null point (previously determined) to the point giving the above retardation is divided by the number of turns per wave to give the waves of angular retardation equal to that produced by the pulse. As noted in Eq. (1), the angular retardation is proportional to the square of the field strength, the constant of proportionality being in turn proportional to the Kerr constant. Hence, a plot of the retardation *versus* the field strength squared will result in a straight line through the origin when everything is properly aligned. If such a graph be obtained for two materials, one of which is CS_2 of known Kerr constant and the other the material in question, the desired Kerr constant will be

$$K_x = K_{CS_2}(\delta_x/F^2/\delta_{CS_2}/F^2).$$

This procedure, of course, presumes the use of the same Kerr cell for both materials.

It should be emphasized that in the determination of the retardation by this method the compensator retardation was varied until the signal on the oscilloscope screen inverted an equal and opposite amount. It must be remembered that the photo-tube amplifier circuit does not interpret d.c. signals; hence, the light at the inverted point is striking the photo-tube at all times except during the pulse. This very fact handicaps this method of measurement for the photo-tube fatigues under conditions of constant light as well as increases the noise level. Both of these effects are clearly visible and introduce appreciable errors.

If a second Kerr cell is used in addition to the compensator, the matter of photo-tube fatigue may be minimized. The plates of the Kerr cells must be so oriented as to give retardations of opposite sign (90° rotation if the Kerr constants are of like sign, 180°, if of unlike sign). Using synchronized pulses on the two Kerr cell plates, the voltages on the cells may be adjusted to nullify each other. Figure 8 is an oscillogram of such a null point with CS_2 and $CHCl_3$. Only the leading and trailing edges are apparent above the noise level of the photo-tube. The Kerr constant for an unknown relative to CS_2 will then be

$$K_{CS_2}{}' = C \cdot K_{CHCl_3}{}' \left(\frac{V_{CHCl_3}}{V_{CS_2}}\right)^2,$$

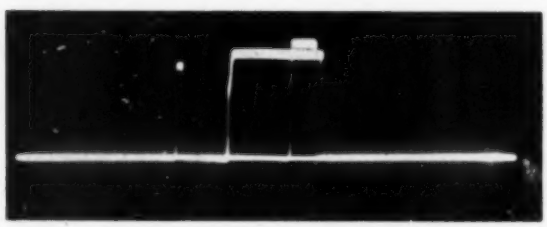


FIG. 6. Oscillogram of optical pulse with nitrobenzene with weak yield at three pulse widths 10, 40, and 150 μsec. 60 p.p.s. 1900 volts, glass separator.

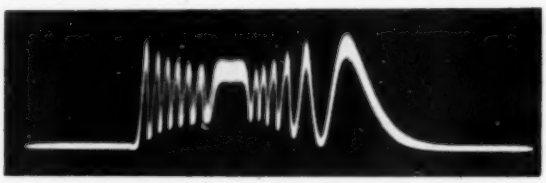


FIG. 7. Oscillogram of nitrobenzene response to high field strength, 28-μsec. width, 400 p.p.s. amplitude 3180 volts with mica spacer, light of 4500A.

where C is a constant accounting for differences in cell spacing and length. Figure 9 shows a graph of the Kerr constant *versus* temperature for CS_2 determined in this manner, with $CHCl_3$ being used as a standard for null purposes. Changes in temperature (t') of the $CHCl_3$ must be accounted for accurately. The agreement with the literature is good.[13] A small amount of constant retardation was introduced by the compensator. This increases the sensitivity of the system without changing the null voltage ratio. Differentiation of Eq. (1) will show the reason for this.[14] No filter was used in determining the curve in Fig. 9, but the data will apply to the average wave-length of approximately 500 mu.

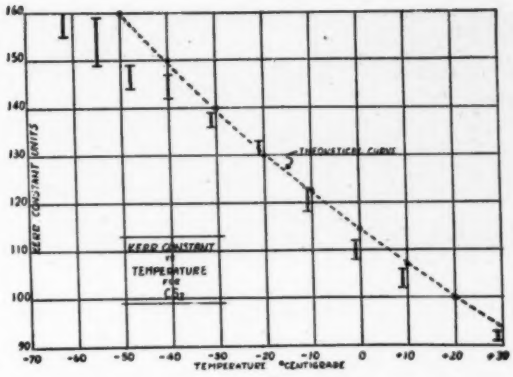


FIG. 9. Kerr constant *vs.* temperature for carbon disulfide.

APPLICATIONS

An application which was mentioned previously as of importance is the study of relaxation in polymers. The molecular orientation hypothesis for the explanation of the Kerr effect as well as usual dielectric absorption effects has been fairly well established;[15] hence, we may picture the Kerr effect as arising from the orientation of molecules or portions of molecules upon the application of an external electrical field. If the molecule is large, and in a viscous medium, the time of orientation may be large. The apparatus described here should be well suited for the observation of this relaxation phenomenon being, in a sense, more direct than the conventional dielectric methods.

Figure 10 is an oscillogram of this effect in Santolite MS, a highly viscous plasticizer, taken at 56°C. At this temperature the molecules or dipoles completely orient in approximately 40 microseconds. However, the relaxation seems to proceed more rapidly than the orientation—an observation which was not expected. At a slightly higher temperature (65°C) the molecules are capable of orienting in 5–10 microseconds as can be seen in Fig. 11. Upon further increase in temperature the orientation becomes more rapid, but the amplitude of the effect (Kerr constant) decreases as one should expect from Eq. (2). If the temperature is lowered and the pulse width increased, it is possible to observe both the orientation and distortion polarizations simultaneously. Figure 12 shows this effect at room temperature with an 800-microsecond pulse. The increased noise level is a result of high amplifier gain. The optical pulse rises very rapidly to a value considerably less than in Fig. 11, then slowly increases until the end of the pulse when it immediately returns to an intermediate level and tapers off slowly to zero. The distortion polarization, of course, accounts for the sharp rise and decline, while the slow change may be attributed to the orientation effect. By such methods it should be possible to isolate and measure the theta-terms of Eq. (2). The dis-

FIG. 8. Typical null point using two-cell technique chloroform and carbon disulfide 10-μsec. width, 300 p.p.s., room temperature.

FIG. 10. Relaxation in Santolite MS 40-μsec. pulse, 300 p.p.s. 3900 volts, temperature 56°C.

[13] A. A. Zuchlka and L. R. Ingersol, J. Opt. Soc. Am. **27**, 314 (1937).
[14] E. F. Kingsbury, see reference 5.
[15] J. W. Beams, see reference 2.

FIG. 11. Relaxation in Santolite MS 20-μsec. pulse, 300 p.p.s. 4000 volts, temperature 63.5°C.

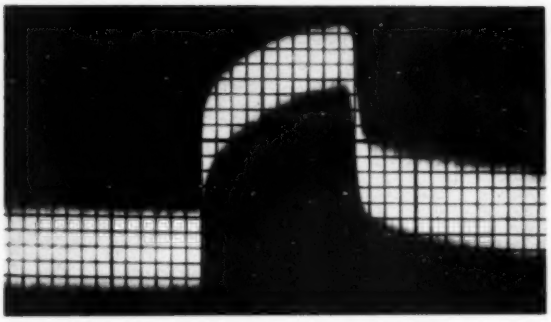


FIG. 12. Oscillogram of Santolite MS at 15°C, 800-μsec. width, 60 p.p.s., 4150 volts; note immediate distortion polarization followed by gradual orientation.

persion region, measured by conventional dielectric methods, lies in the same range of temperature for the conventional dielectric method as for the electro-optic, giving strong indication to the fact that the same unit (molecule or dipole) is orienting in both cases.

This relaxation effect in polymers can be observed dielectrically with pulsed fields by placing a test condenser of the dielectric directly across the load resistor of the pulse generator cathode-follower circuit and observing the relaxation with a cathode-ray tube. The effect is not as pronounced, because of other instrumental effects, but has been observed to be temperature and time dependent in the same manner as in the Kerr effect.

Further studies are in progress correlating this Kerr effect with conventional dielectric properties and viscosity for a series of primary alcohols.

On the Localized Afterglows Observed with Long Sparks in Various Gases

J. D. Craggs and W. Hopwood
Research Department, Metropolitan-Vickers Electrical Company, Ltd., Manchester, England

and

J. M. Meek
Department of Electrical Engineering, Liverpool University, Liverpool, England

(Received June 2, 1947)

Simultaneous photographs, taken in two perpendicular directions, have been obtained of long sparks in various gases at 1-atmosphere pressure. The experiments were undertaken in order to decide whether the localized afterglows, noticed as long streaks in rotating mirror spark photographs, were caused by some discharge mechanism or by purely geometrical effects resulting from the tortuous nature of the spark channels. The results show that the second explanation is not feasible. Some qualitative interpretations of the effects are given.

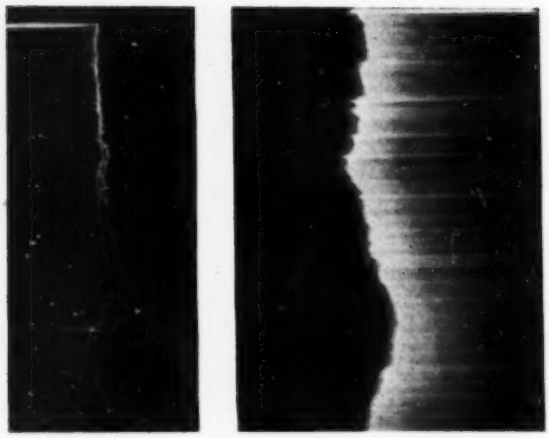
I. INTRODUCTION

THE study of spark discharges in gases has usually been confined to the examination of such discharges in air and to the earlier parts of the process, e.g., the formation of streamers.[1] Recently the later phases of spark formation (spark channels) have been re-examined and some of their properties have been described by Raiskiy (1940),[2] Flowers (1943),[3] and by Craggs and Meek (1946),[4] following the earlier work of several authors such as Lawrence and Dunnington (1930),[5] Ollendorff (1933),[6] and others. Such spark channels may be considered, in brief, to consist of a highly conducting path (i.e., of low voltage gradient) in which the electron temperature may be 10,000–20,000°, and approximating in some aspects to a nascent arc of artificially restricted diameter. Rotating mirror camera records of lightning, taken mainly by Schonland and his collaborators to whom much of the best work is due, and of laboratory sparks (Allibone and Meek (1938), (a) and (b),[7,8] and others) have shown, however, that the last stages of a spark discharge, perhaps even when the current flow has sensibly ceased (Craggs and Meek)[4] are represented by localized afterglows. The significance of the latter has often been overlooked or minimized, and the purpose of the present paper is to present an account of further preliminary experiments made on this effect.

II. DISCUSSION OF PREVIOUS EXPERIMENTS

Many previous papers, e.g., those of Allibone and Meek[7,8] tend to ignore the afterglow streaks, shown clearly in a typical record such as Fig. 1 (b)(ii) of Allibone and Meek's second paper[8] and concentrate on other aspects of the discharge. Walter (1935),[9] in a discussion on Hoffert's early photograph (1889)[10] of a multiple lightning stroke, noticed the streaks and stated that "they are by no means due to an afterglow of the lightning channel produced by thermal or phosphorescent causes, but they are always associated with a real after-discharge in the channel, i.e., with an electric current following the main discharge along the same track." Walter considered that this conclusion is justified by the fact that the streaks often increase in brightness at an appreciable time after the light from the main channel has greatly decayed. An excellent example of this effect has been published by Allibone and Meek[8] (Fig. 1 (a)(iii)); in such cases there is almost certainly an appreciable oscillatory flow of current, but such currents should be constant at all points along the dis-

[1] L. B. Loeb and J. M. Meek, *The Mechanism of the Electric Spark* (Stanford University Press, Stanford University, California, 1941).
[2] S. M. Raiskiy, J. Tech. Phys. U.S.S.R. **10**, 529 (1940)
[3] J. W. Flowers, Phys. Rev. **64**, 225 (1943).
[4] J. D. Craggs and J. M. Meek, Proc. Roy. Soc. **A186**, 241 (1946).
[5] E. O. Lawrence and F. G. Dunnington, Phys. Rev. **35**, 396 (1930).
[6] F. Ollendorff, Archiv f. Elektrotechnik. **27**, 169 (1933).
[7] T. E. Allibone and J. M. Meek, Proc. Roy. Soc. **A166**, 97 (1938).
[8] T. E. Allibone and J. M. Meek, Proc. Roy. Soc. **A169**, 246 (1938).
[9] B. Walter, Phil. Mag. **20**, 1144 (1935).
[10] H. H. Hoffert, Phil. Mag. **28**, 106 (1889).



(a) (b)

FIG. 1. Rotating film photographs of the point/plane sparks in air (1-atmos. pressure); (a) had 100-cm gap and a high circuit resistance; (b) had 130-cm gap and a low circuit resistance and is an enlarged record.

charge, a supposition which was apparently overlooked by Walter and which unless other effects are present would demand uniform afterglows along the channel at any particular instant. It is therefore very unlikely that "follow currents" can account for the localized streaks of afterglow, although accentuation of the latter by the former probably occurs.

McEachron and McMorris[11] appear to accept Walter's view in a modified form and suggest that very small (<100 amp.) follow currents may be responsible, but again ignore the peculiar localized nature of the afterglows. It is appropriate to state here that the measurement of small follow currents is difficult, and their existence in many laboratory experiments cannot be denied. Such measurements are now projected in this laboratory, although the arguments to be advanced below tend, it is suggested, to minimize their fundamental importance. The lightning photographs of Malan and Collens,[12] especially their Figs. 13–17 include some excellent records, and those authors agree with the conclusion of McEachron and McMorris.

Flowers[13] subscribes to an alternative theory of the afterglow streaks and states that "portions of the path (of the discharge) which are projected parallel to the direction of motion of the film or the lens provide a greatly increased exposure time and proportionate reduction in resolution. Path components which lie along the photographic axis increase the light intensity. Both effects produce excessive exposures and are associated with the streaks that are observed in nearly all Boys camera photographs." It is clear that, whilst Flowers does not apparently in this communication attribute the streaks directly or wholly to the effects described, which he states "are associated" with them, the effects of bends could contribute to or produce highly localized streaks. This explanation had occurred independently to us some years ago, but was immediately rejected, since the length of the afterglows seem to be of quite a different order of length from the average effective length of a bend in the channel.[14] It appeared necessary to disprove this theory, if possible, before embarking on further experiments since, if true, the bend

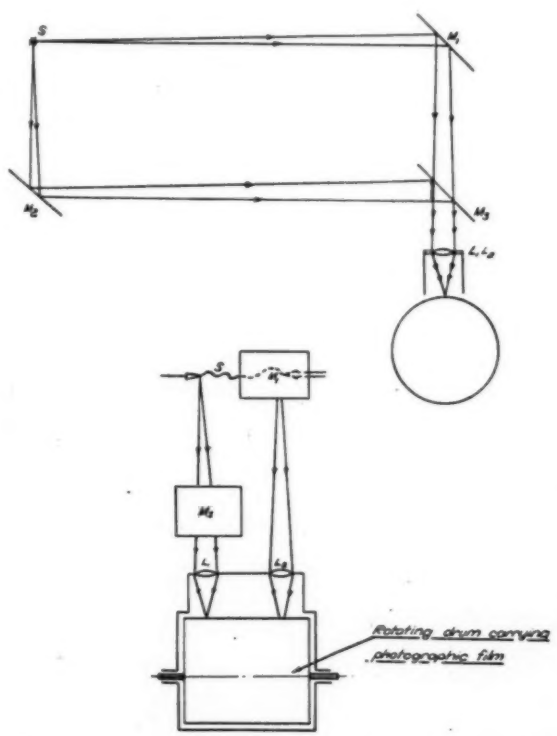


FIG. 2. Optical system used in right-angle spark photography; M_1, M_2, and M_3 are mirrors; S is the spark discharge; L_1 and L_2 are lenses.

[11] K. B. McEachron and W. A. McMorris, Gen. Elec. Rev. **39**, 487 (1936).
[12] D. J. Malan and H. Collens, Proc. Roy. Soc. **A162**, 175 (1937).
[13] J. W. Flowers, Gen. Elec. Rev. **47**, 9 (1944).
[14] J. D. Craggs, Quart. J. Roy. Met. Soc. **70**, 167 (1944) (discussion on paper by T. E. Allibone).

theory would entirely remove the physical importance of the effect. It is important to note that Flowers' suggestion refers to bends in the main channel and not in the much dimmer and more transitory leader strokes. (It does in fact appear from many records (e.g., those of Fig. 1) that the streaks arise in the leader strokes or streamers and persist through the later stages of the spark channel's formation and extinction.)

In the discussion following a paper by Allibone (1938),[15] O'Dea made the suggestion that the streaks may be due to the persistence of "ionized or ozonized pockets." Allibone, in reply, referred to the theoretical work of Neugebauer[16] on ball lightning. The present authors favor O'Dea's proposition, with the important correction that the presence of free electrons and ions, demanded by Neugebauer's theory, may not be necessary since the afterglows may be explained on other grounds. (Craggs and Meek,[4] with references to earlier work from 1943 by the same authors.)

As mentioned above, it appeared necessary first to investigate the possibility of a geometrical explanation of the afterglow streaks, i.e., the effect of sustained illumination of the photographic film caused by the increased exposures at channel bends, either in or perpendicular to the plane of the film. This paper describes such experiments.

III. THE EXPERIMENTAL METHOD

The most obvious method for assessing the importance of the geometrical bend theory, outlined above, is to photograph the sparks with a rotating film or mirror camera in two different, preferably perpendicular, directions at once. If the results should, for example, show that the luminous afterglow streaks are noticed at points where no channel bending is perceptible, then it would appear reasonable to discount the bend theory and to suppose that some effect in the ionized or excited gas in the spark channel is operative. Further investigations to elucidate the nature of the effect would then be warranted.

Strigel[17] is apparently the only author to have photographed sparks previously in this way, but his object was not to investigate the afterglow streaks, which are indeed not mentioned in his paper, and it is not possible to use Strigel's reproductions of spark photographs in deciding the point at issue in the present paper.

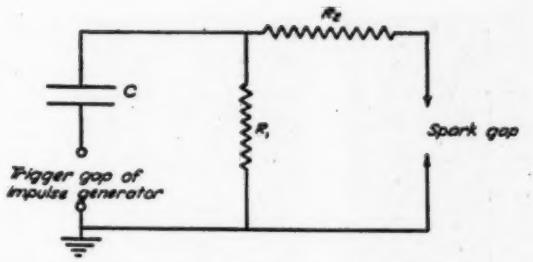


FIG. 3. Impulse generator; $C=6$ series stages of 0.22 uf per stage. Total discharge capacity $=0.037$ uf; $R_2=210$ ohms; $R_1=35{,}000$ ohms; charging voltage $=70$ kv.

The optical system is shown in Fig. 2, in which the rotating drum (3000 r.p.m.) carries the photographic film. The drum diameter was approximately 12". The records (see Fig. 5) were therefore composed of pairs of photographs arranged one above the other on the film, each pair corresponding to one spark.

The spark was produced with a 2-million volt impulse generator (0.11 mfd/stage) rearranged to give effectively the circuit of Fig. 3, in which a peak current of 2000 amp. was used. The point/point gap of 20 cm needed some 280 kv for breakdown in air. Other experiments were made at 1000 and 500 amp., with suitable alterations to the circuit of Fig. 3. A current oscillogram for 2000-amp. peak current is given in Fig. 4.

For gases other than air, a glass vacuum chamber (30-cm diameter, 60 cm long) was used, pumped down before gas admission to a pressure of about 1 mm Hg. The gases from cylinders were then introduced to a pressure of 80 cm Hg, in order to prevent contamination by air leakage. It was found to be necessary to attach a black cloth inside the glass cylinder to prevent reflections. The thickness of the glass chamber was about 1 cm so that ultraviolet radiations from the sparks were prevented from reaching the film.

About 30 films, each with 10–20 pairs of records, were taken.

IV. RESULTS OF THE EXPERIMENTS

Figure 5 shows some results (full-size reproductions of original negatives), obtained with 2000-amp. discharges in oxygen (3 sparks), hydro-

[15] T. E. Allibone, J. Inst. Elec. Eng. **82**, 487 (1938).
[16] T. Neugebauer, Zeits. f. Physik **106**, 474 (1937).
[17] R. Strigel, Wiss. Veröff. aus den Siemens Werken. **15**, Part 3, 13 (1936).

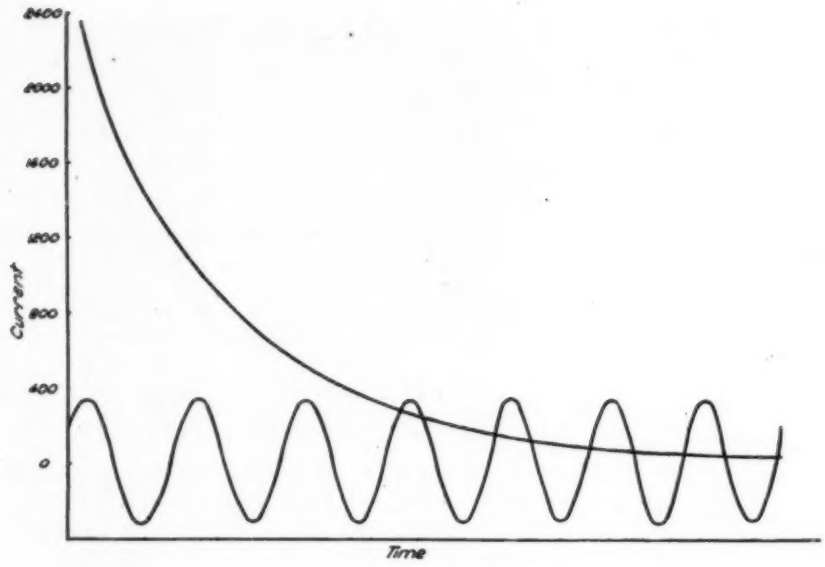


FIG. 4. 2000-amp. current oscillogram for spark discharge in argon. The timing oscillation has a period of 5 microsecs.

gen (2 sparks), and nitrogen (2 sparks) which are typical of the many other photographs taken (1 cm on the time axis≡240 microseconds). The value of the method for assessing the importance of channel bends is at once apparent. Oxygen and nitrogen show pronounced afterglow streaks which are highly localized, particularly in nitrogen. The formation of branched channels was found with oxygen and nitrogen but not with hydrogen, and the latter gas showed very short afterglows, in agreement with previous experiments. The spark channels in hydrogen have a pronounced tendency to form bends which are particularly sudden and well defined. The very short afterglows in hydrogen are fairly localized, as shown by the tufted appearance of the channels on their afterglow side.

It is clear, especially from the oxygen records of Fig. 5 (a) that the streaks are formed in both records at the same points in the channel, not necessarily where bends or inflections are noticed, and are generally of the same length. These findings are in general disagreement with the geometrical theory. Where channel division occurs (Fig. 5 (a)) the different views of the spark may be different largely because of the geometrical (not time) separation of the channel branches with subsequent overlapping and confusion of the afterglows. The exceptionally tortuous bends in hydrogens do not give enhanced streaks, as would be expected from the geometrical theory which does not invoke any difference in channel structure from gas to gas.

When the argon results, typical specimens of which are given in Fig. 6, are examined it seems even clearer that the afterglow streaks are associated with localized discharge conditions in the channel. Figure 6 (b) particularly demonstrates the evidence; argon spark channels show only a tendency to form bends which are, even then, very gradual and yet the streaks are particularly long and well defined, and situated clearly at the same points and for the same times, in both records for each spark. Figure 6 (a) shows an exceptionally straight spark. Figure 6 (b) demonstrates another effect noticed in argon, i.e., the formation, at approximately the same point in the channel even for different sparks, of a single well-defined streak. This phenomenon is probably, it is thought, connected with the earlier parts of the discharge process and if substantiated this conclusion would constitute irrefutable evidence against the supposed importance of channel bends. It is, however, considered that the data given in Figs. 5 and 6 and described above constitute sufficient proof of the genuine physical basis of the afterglow streaks which, while not being explained finally in this paper (see below), assume a greater physical interest than they would have possessed if explained in the manner suggested by Flowers and others.

The use of pairs of photographs, such as those of Figs. 5–7, enable the true length, in three dimensions, of the spark channels to be found. This may be useful in experiments on the voltage drop in spark discharges, where the true channel length if it differs appreciably from the two dimensional value, should be known. This effect would probably be noticeable only with hydrogen channels; further experiments on these lines are in progress.

Before describing further analyses of records similar to those of Figs. 5 and 6, it is necessary to mention the very frequent and often complex division of channels noticed in argon and demonstrated in Figs. 6 (c), (d), and (e). The other gases so far examined show relatively infrequent branching which generally extends (Figs. 5 (a) and (d)) only for a fraction of the gap length. Argon, on the other hand, often shows (Figs. 6 (c) and (d)) complete branching of the gap by

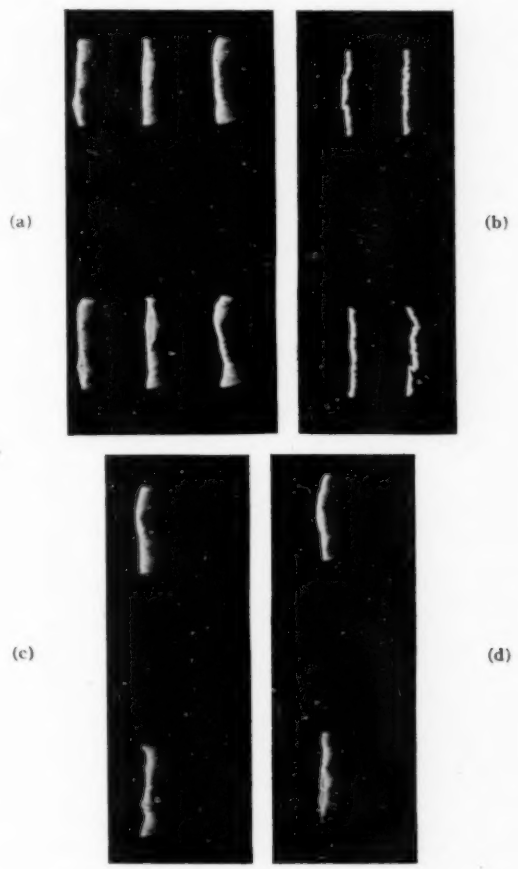


FIG. 5. Photographs of sparks in various gases; (a) oxygen, (b) hydrogen, (c) and (d) nitrogen.

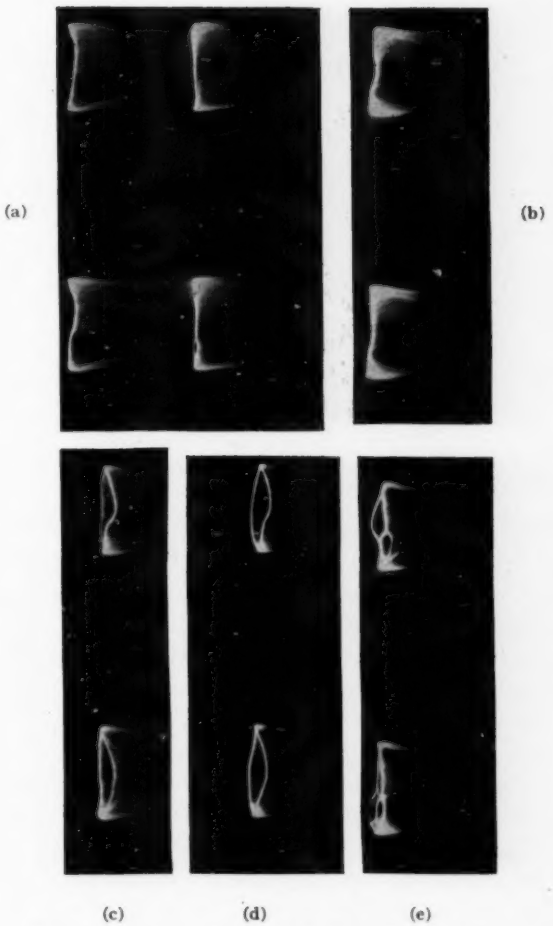


FIG. 6. Photographs of spark discharges in argon.

divided channels and also multiple division, a complex case of which is shown in Fig. 6 (e). The value of bi-directional photography in the elucidation of such phenomena is obvious. Yet another interesting effect is seen in Fig. 6 (c) (and many other records) where the afterglow streaks are seen to change their positions in space, as time advances. It seems probable in Fig. 6 (c) that the right-hand channel of the upper photo-

TABLE I.

Gas	Total No. of sparks examined	No. of sparks with two channels	No. of sparks with three or more channels
Hydrogen	27	0	0
Oxygen	43	27	5
Nitrogen	39	22	1
Argon	85	35	6
Air	71	10	0
Methane	20	1	0

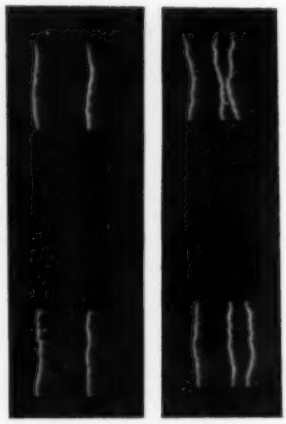


FIG. 7. Photographs of spark discharges in methane.

graph is the left-hand channel of the lower photograph, since this branch shows, in both photographs, relatively little afterglow. These effects are being investigated further.

Table I gives some statistical data on channel division, which are discussed below.

The methane photographs (Fig. 7) were taken in the hope that the known high absorption of photo-ionizing radiations in such a complex molecular gas[18] would lead to the demonstration of unusual properties of the spark channel. The peculiar rippling structure of the methane channels are indeed noticeable, particularly on the leading edge of the channels where the afterglows, being absent, do not cause confusion. Such tortuous channels are to be expected, on general grounds, if the photo-ionization process at the streamer tip is inefficient.[1]

Microphotometer tracings of typical records in several gases were taken at right angles to the mean channel direction along afterglow streaks and across parts of the channel where such streaks were absent. Figures 8 (a)–(c) show, respectively, tracings for argon, oxygen, and nitrogen. The right-hand tracings for each gas were taken along an afterglow. Figures 9 (a) and (b) show hydrogen records, the former taken in the afterglow. 1 cm on the time axis (right to

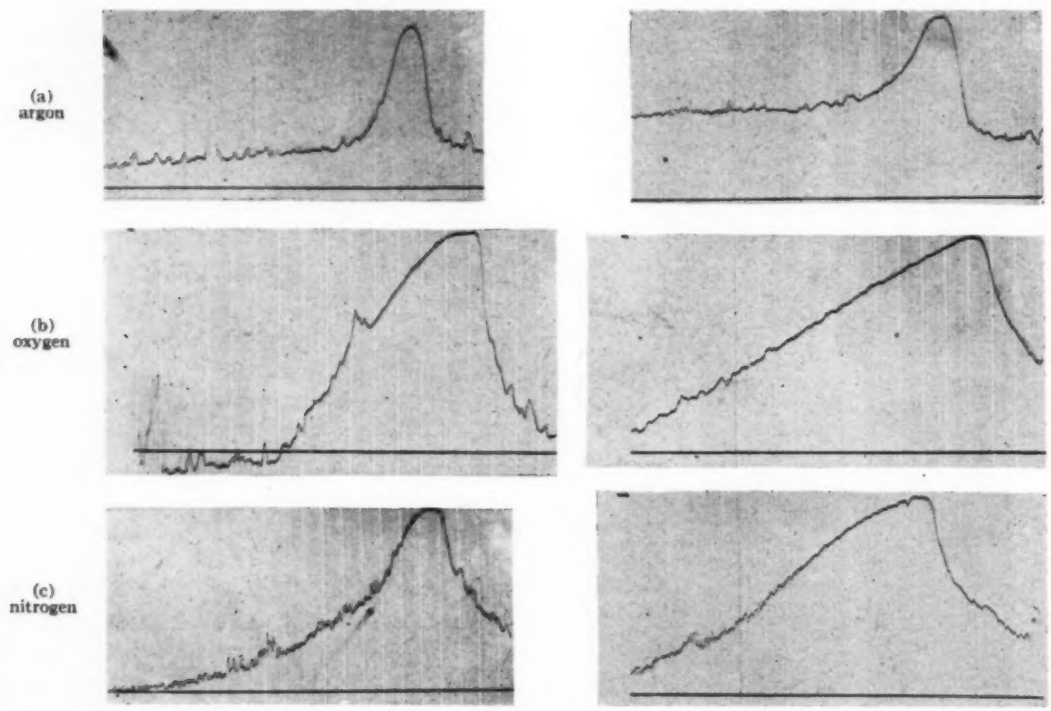


FIG. 8. Microphotometer tracings of spark photographs, such as those of Figs. 5–7. Time scale is from right to left. The left-hand records are for normal parts of the channels, and the right-hand records are for points in the channels which show the localized afterglows.

[18] S. A. Korff and R. D. Present, Phys. Rev. **65**, 274 (1944).

left) of the figures represents 21 microseconds. It is possible to deduce from these records the light emission/time relation for any part of the channel if the response curve of the film is known. The latter quantity was determined by exposing samples of the film to 2000-amp. sparks, of the type studied in this paper, through an Ilford stepped filter (neutral) which was in turn calibrated with the usual photo-cell arrangement. Exposures were also made to steady light. A typical calibration record, using spark light, is shown in Fig. 10. Figures 11 and 12 show the results of the analysis, i.e., the gamma-curves for the film using spark and steady light, respectively.

Since the light from the spark channels is relatively very intense, with respect to the afterglows, the calibration records, such as Fig. 10, show the film's response to the channel and not the afterglow light. However, since Fig. 11 shows that the difference in the responses to channel and steady light is negligible and the response to the afterglow light must lie between the limits

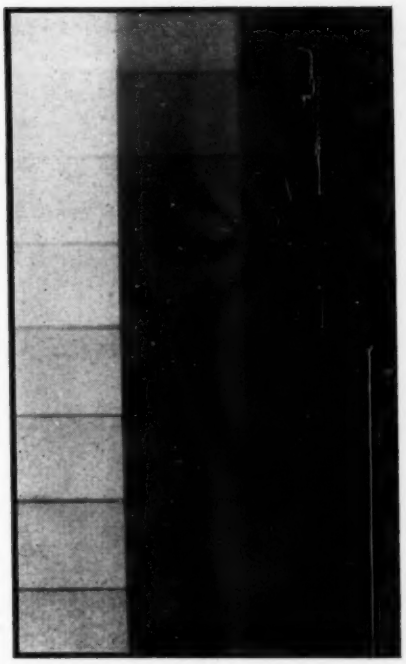


FIG. 10. Step wedge photograph with argon spark light, for film calibration purposes.

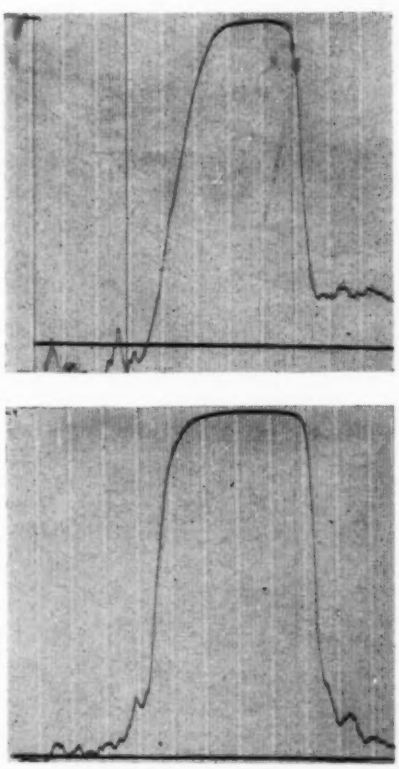


FIG. 9. Microphotometer tracings of hydrogen spark photographs (see Fig. (5b)); time scale is from right to left.

set by Fig. 11, it is reasonable to take the spark-light data of Fig. 11 as referring to the afterglow light.

Figure 12 shows the final light intensity/time curve for an argon afterglow, with traces taken along an afterglow streak and across the channel at a point where no appreciable streak was observed.

V. DISCUSSION OF THE RESULTS

It seems clear from the data given in Section IV that the suggestion of the bends being responsible, in a geometrical sense, for the afterglow streaks is untenable. The influence of the magnetic field of the discharge channel on the afterglow should be considered. This effect was suggested originally by Anderson and Smith[19] to account for similar localized afterglows, described as being due to vapor jets, with exploded wires, and it was pointed out that the magnetic field tended to prevent charged particles from leaving the conducting channel especially at the points of strongest fields. The field is weakest, however, at the outside of a sharp bend where afterglow

[19] J. A. Anderson and S. Smith, Astrophys. J. **64**, 295 (1926).

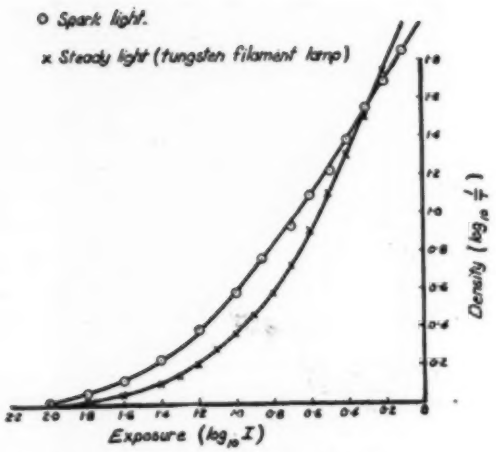


FIG. 11. Calibration of the photographic films with the step wedge.

streaks are often, but not always, encountered; and it is clear that no definite conclusion as to the relevance of this effect which could, at least, explain the localization of the afterglow can be made in the absence of further data.

The following mechanisms could account for general, i.e., non-localized afterglows. The influence of the low current still passing through the discharge channel after say 100 microseconds is not so easily dismissed. Figure 4, which is a reproduction of a current oscillogram for argon (no apparent systematic difference in residual currents for the different gases was noticed, but the accuracy of estimation of these low currents by the present method is not great) shows that the current at 100 microseconds from the beginning of current flow is about 20 amp. The records showed currents varying from this order of magnitude to almost zero. However, the passage of a small residual current cannot by itself account for the localized nature of the afterglows, since the current must also flow through the relatively dark regions of the discharge. It has been shown[4] that the afterglows persist even when square-wave pulses of current are used, and it therefore seems reasonable to postulate the existence of afterglows more or less independent of even a small current flow in the gap.

Afterglows in low pressure discharges are well known, by the pioneer work of Rayleigh, and later, many others on active nitrogen. Kenty[20] has observed an afterglow in argon, following a low pressure arc discharge, and has attributed this to electron/ion recombination. Afterglows in luminous clouds, following arc discharges at atmospheric pressure, have been noticed by O'Doherty[21] and others.

The afterglows in hydrogen have been discussed elsewhere, and it is sufficient here to state that they are probably due to electron/ion recombination. The afterglows in argon (and certain other gases, e.g., oxygen and nitrogen) could presumably be due either to recombination of various kinds (electron/ion mechanism in argon, where attachment is small) or to the persistence of metastable atoms. It is hoped that further experiments by the authors will clarify the matter.

It seems unlikely that persistence of metastable atoms in argon is the sole cause of the afterglows, since addition of oxygen or hydrogen to argon in reasonable quantities has only a minor effect on them,[4] although both diluents in very small amounts have a great effect on the dielectric recovery of argon spark gaps. If, therefore, electron/ion recombination is the main cause of the visible afterglow streaks it is of interest to consider why the afterglows in hydrogen and argon should be widely different in duration. The recombination coefficients themselves are not likely to be widely different, and the problem would then resolve into a study of the energy losses (by elastic collisions) of the electrons in the afterglow, since the recombination cross sections increase with decreasing electron energy. It is apparent at once from the known data on the Ramsauer effect etc., that for equal electron energy, the elastic collision cross sections are much greater in hydrogen than in argon for low electron energies, so that the electrons would attain thermal energies more quickly in the former gas and would therefore be removed by recombination more quickly. Unfortunately, Ramsauer data are not available for electron temperatures of ~500–1000° (mean energy ~0.1 volt or less), which probably obtain in the afterglow conditions.[22] Cravath[23] has worked out the average energy loss (f) per encounter, in elastic

[20] C. Kenty, Phys. Rev. **32**, 624 (1928).

[21] J. J. O'Doherty, Nature **153**, 558 (1944).
[22] J. D. Craggs and W. Hopwood, in the press.
[23] A. M. Cravath, Phys. Rev. **36**, 248 (1930).

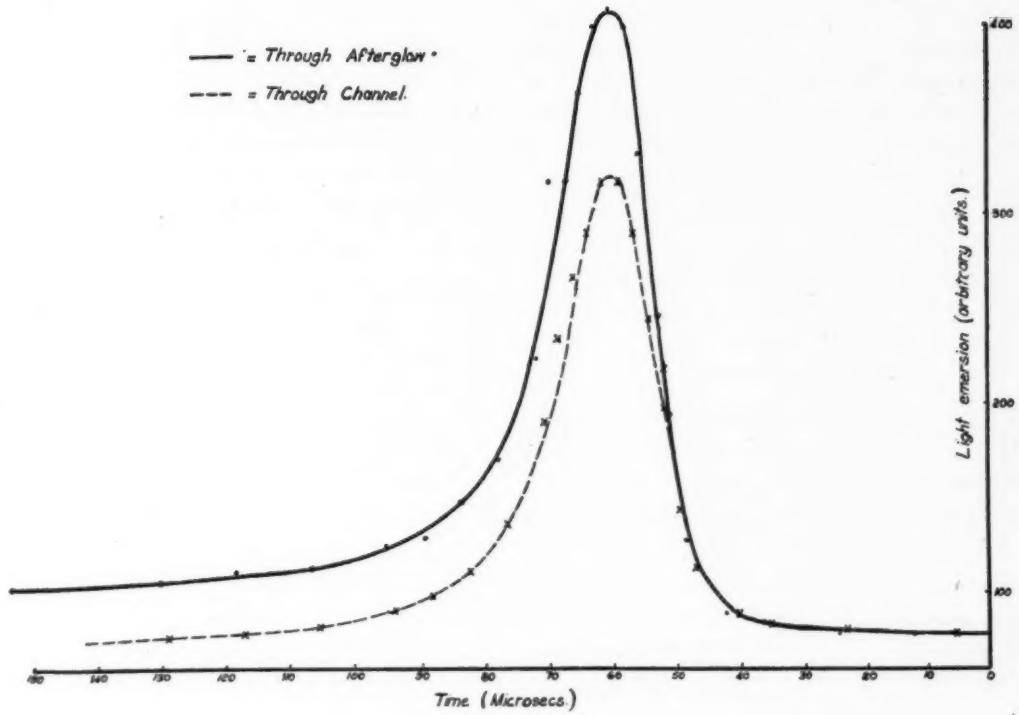


FIG. 12.

collision, expressed as a fraction of the average energy of the electron:

$$f=\frac{8}{3}\frac{mM}{(m+M)^2}\left(1-\frac{T_M}{T_e}\right),$$

where T_M is the equivalent temperature of the molecules present and T_e is the equivalent temperature of the electrons. Generally, $T_M \ll T_e$ so that

$$f\simeq\frac{8}{3}\frac{mM}{(m+M)^2}$$

($m+M$ are the electronic and molecular masses, respectively). It can be shown that f is very great until $T_e \sim T_M$. Using elastic cross sections (P_e) given by Normand[24] and Brode,[25] the data of Table II are easily calculable. In some cases (at the lowest temperatures) extrapolation of the elastic cross-section/electron energy curves has been necessary; this is clearly a most unsatisfactory procedure, but it is the only possible one in the circumstances and leads only to approximate results. T_M has been taken at 300°K. If a higher value for T_M is permissible, then f will be lower than the respective values in Table II, particularly at $T_e \sim 500$°K. The procedure for determining the total f is (i) calculate the mean electronic energy and velocity (v) from T_e, which is assumed to correspond to a Maxwellian distribution; (ii) determine the mean free path λ for elastic collisions, i.e., $1/P_e$, for the known electronic energy; (iii) from λ and v the total number of collisions in a given time, for each one of which a fractional energy loss f is suffered, can be found, and so (iv) the total f for that time can be computed.

TABLE II.

Gas	Electron temp. °K	Total f for time of 10^{-8} sec.
Hydrogen	10,000	32
Hydrogen	1,000	39
Hydrogen	500	20
Argon	10,000	0.093
	5,000	0.066
	1,000	0.062

It must be emphasized again that the results are only approximate, but they show that the energy loss for electrons in hydrogen in our par-

[24] C. E. Normand, Phys. Rev. **35**, 1217 (1930).
[25] R. B. Brode, Rev. Mod. Phys. **5**, 257 (1933).

ticular conditions is far higher than for argon, and so a recombination spectrum would persist for a longer time (other things being equal) in argon than in hydrogen. The important influence of T_M at values of $T_e \sim T_M$, and our ignorance of its magnitude, render further speculation unprofitable. Experiments now being made will, it is hoped, provide more accurate data on the values of T_M and T_e obtaining in the decaying channels of spark discharges.

VI. CONCLUSIONS

1. Photographs have been taken with a rotating film camera, and simultaneously in two perpendicular directions of spark discharges.

2. The photographs show, we believe conclusively, that the localized afterglow streaks seen in records of this kind are not due to bends, in the spark channels, causing artificially long exposures by virtue of the enhanced depth of glowing gas, etc.

3. The causes of the afterglows are not known with certainty in all cases.

4. We believe that further experiments, now in hand, will give further information on localized afterglows in spark channels and probably, therefore, on ball or bead lightning[26] which we consider to be probably due to the same causes.

VII. ACKNOWLEDGMENTS

The authors are grateful to Mr. F. R. Perry for the interest he has shown in, and the support he has given to, the work. They also wish to thank Mr. G. C. Williams and Mr. W. Dale for assistance with the camera and high speed oscillograph, respectively. Finally thanks are expressed to Sir Arthur P. M. Fleming, C.B.E., D.Eng., Director of Research and Education, and to Mr. B. G. Churcher, M.Sc., M.I.E.E., Manager of Research Department, Metropolitan-Vickers Electrical Company, Ltd., for permission to publish this paper.

[26] D. G. Beadle, Nature **137**, 112 (1936).

Erratum: Conductivity of Metallic Surfaces at Microwave Frequencies

[J. App. Phys. **18**, 629 (1947)]

E. MAXWELL
Massachusetts Institute of Technology, Boston, Massachusetts

1. p. 629, line 5 of fourth paragraph: Replace $J_0^2/2\sigma\delta$ by $(J_0/2)^2(\delta/\sigma)$.

2. p. 629, line 8 of fourth paragraph: Replace J_0 by $J_0/\sqrt{2}$.

3. p. 629, second column, equation should read:

$$\int\int\left(\frac{J_0}{2}\right)^2\frac{\delta}{\sigma}ds = \int\int\frac{H_t^2}{2\sigma\delta}ds.$$

4. p. 632, Fig. 3b should be inverted.

5. p. 634, line 21 should be changed to read "a and c are incommensurable."

6. p. 636, line 25, replace "conductivity" by "resistivity."

Letters to the Editor

Quantitative Analysis of Mixed Powders with the Geiger-Counter X-Ray Spectrometer

ZIGMOND W. WILCHINSKY
Esso Laboratories, Standard Oil Company of New Jersey, Baton Rouge, Louisiana
July 23, 1947

A USEFUL characteristic of a Geiger-counter spectrometer is its ability to measure diffracted x-ray intensities for several samples in succession, on the same intensity scale. This property greatly facilitates the quantitative determination of crystalline components in a sample of mixed powders. In particular, it becomes quite easy to determine the amount of a crystalline component in the presence of other components which may be crystalline or amorphous, identified or unidentified. A simple expression relating the weight fraction of a crystalline component to experimental quantities is developed below.

For a flat sample used in a North American Philips x-ray spectrometer,[1] the relationship between the sample and the incident and diffracted x-ray beams is indicated in Fig. 1. Let us assume that simple x-ray diffraction

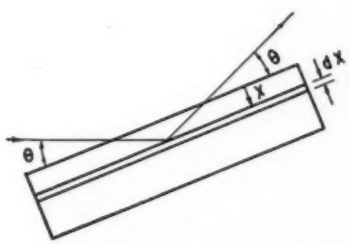


FIG. 1. Diffraction of x-rays from a flat sample in a Geiger-counter x-ray spectrometer.

theory is applicable. Then, for a chosen diffraction line of the ith crystalline component, the contribution to the integrated intensity from the elemental layer dx is

$$dP_i=\frac{AJ_i}{\sin\theta}W_i\rho\exp(-2\mu_l x/\sin\theta)dx, \tag{1}$$

where A = effective cross-section area of the beam at the sample, J_i = contribution to the integrated intensity per unit weight of sample on the basis of no absorption, W_i = weight fraction of the ith component in the sample, ρ = mean density of the sample, μ_l = linear absorption coefficient for the sample. For a sample of sufficient thickness to be considered opaque, P_i is obtained by integrating the right-hand side of (1) from $x=0$ to $x=\infty$, the result being

$$P_i=\tfrac{1}{2}AJ_iW_i\rho/\mu_l=\tfrac{1}{2}AJ_iW_i/\mu, \tag{2}$$

(2) where $\mu=\mu_l/\rho$ is the mass absorption coefficient for the sample. The factor AJ_i can be evaluated by considering a sample consisting only of the ith component; the integrated intensity, P_{0i}, for this case, by Eq. (2) is

$$P_{0i}=\tfrac{1}{2}AJ_i/\mu_i, \tag{3}$$

where μ_i is the mass absorption coefficient for the ith component. Eliminating AJ_i from (2), one obtains the useful and very simple expression:

$$W_i=\frac{P_i}{P_{0i}}\frac{\mu}{\mu_i}. \tag{4}$$

The value of μ_i can be readily computed from the chemical formula of the ith component and a table of mass absorption coefficients for the elements. Similarly, the value of μ can be computed from an elemental analysis for the sample.

[1] H. Friedman, Electronics **18**, 132 (1945).

The Conversion Factor for *kX* Units to Angström Units

ELIZABETH ARMSTRONG WOOD
Bell Telephone Laboratories, Murray Hill, New Jersey
July 21, 1947

X-RAY wave-lengths have been expressed in X units. The X unit is defined in terms of the calcite spacing and is nearly 10^{-11} cm, but is now known to differ from 10^{-11} cm by about 0.2 percent. During the last twenty-five years x-ray diffraction workers have expressed x-ray wave-lengths and crystal dimensions in terms of a unit which was 1000 X units, but instead of calling it 1000 X units have erroneously called it an Angström unit. In recent years, the x-ray diffraction groups have agreed to use the term kilo X unit (abbreviated kX) in place of the incorrectly used Angström unit, until agreement was reached on the best conversion factor to use for converting from kX to Angström units. Agreement on the factor has now been reached.

As Secretary of the American Society for X-Ray and Electron Diffraction, I have been instructed to call the attention of American x-ray workers to the following announcement which appeared in the January 1947, issue of the Journal of Scientific Instruments. Because of its importance it is here reproduced in its entirety.

At the annual conference of the X-ray Analysis Group of the Institute of Physics in July 1946 it was announced that agreement had been reached concerning the factor for converting measurements in kX units to Angström units. The factor agreed upon, after consultation with the American Society for X-ray and Electron Diffraction and Prof. Siegbahn was 1.00202. This factor is probably correct to 0.003 percent. Since wave-lengths in X-units have been measured to an accuracy of 0.001 percent, the wave-lengths in Angström units can be taken as accurate to 0.004 percent in general.

The following is a list of values of wave-lengths in Angström units of certain emission lines and absorption edges in common use. The column headed $K\alpha$ gives the mean value of $K\alpha_1$ and $K\alpha_2$, $K\alpha_1$ being allowed twice the weight of $K\alpha_2$.

Current values of the physical constants, such as those quoted by Birge in the 1941 volume of the Physical Society's *Reports on Progress in Physics*, should be used in conjunction with these wave-lengths. In particular density ρ is given by the equation

$$\rho = 1.66020 \Sigma A / V$$

where ΣA is the sum of the atomic weights of the atoms in the unit cell, and V is the volume of the unit cell in A^3.

	$K\alpha_1$	$K\alpha_2$	$K\alpha$	$K\beta_1$	Absorption edge
Cr	2.28962	2.29352	2.2909	2.08479	2.0701
Mn	2.10174	2.10570	2.1031	1.91016	1.8954
Fe	1.93597	1.93991	1.9373	1.75654	1.7429
Co	1.78890	1.79279	1.7902	1.62073	1.6072
Ni	1.65783	1.66168	1.6591	1.50008	1.4869
Cu	1.54050	1.54434	1.5418	1.39217	1.3802
Zn	1.43510	1.43894	1.4364	1.29520	1.2831
Mo	0.70926	0.71354	0.7107	0.63225	0.6197
Rh	0.61326	0.61762	0.6147	0.54559	0.5341
Pd	0.58545	0.58982	0.5869	0.52052	0.5090
Ag	0.55941	0.56381	0.5609	0.49701	0.4855

It is recommended that in any published work the values of the wave-lengths used should be explicitly stated.

W. L. Bragg
Chairman, X-ray Analysis Group of the
Institute of Physics (England).

New Books

Astronomy

By John C. Duncan. Fourth Edition. Pp. 500. Harper and Brothers, New York, 1946.

Professor Duncan's book now in its fourth edition offers an excellent text for the beginning college student of astronomy. This book may be read with profit by anyone who wishes to obtain a knowledge of modern astronomy.

The introductory chapter includes an exceptionally fine set of star charts showing the naked eye stars, though a more detailed description of how to use them would have helped the student. Following the chapter on the celestial sphere and coordinate systems is one on telescopes. Though the Schmidt telescope is described, the reader is not given the simple explanation of the function of the correcting plate. A new and very excellent chapter on Celestial Navigation has been added following the chapter on the Earth and Sky. Chapters on the Orbital Motion of the Earth, the Moon, and Eclipses are followed by a chapter on Spectroscopy which is introduced early in order to make use of the principals involved in the chapter on the Sun which follows. The Paths of the Planets and the Law of Gravitation are next taken in order; the latter seems rather detailed in parts for a book designed for beginning students. Chapters on the Planets, Meteors, and Comets precede a discussion of the evolution of the solar system, a chapter which dwells unduly long on the older theories at the expense of the more modern theories of Spitzer, Littleton, and Russell. However, since even modern theories on this subject are still far from giving us a complete answer to this riddle, perhaps Dr. Duncan is justified in these omissions.

The stars and their motions are next discussed and then follows a completely up-to-the-minute (1945) chapter on Physics of the Stars. Here the student may find "the going a little tough," and will undoubtedly want to review the earlier chapter on Spectroscopy before wrestling with even elementary astrophysics. After a chapter on Variable Stars is one on Star Clusters and Nebulae in which the careful student may be somewhat confused to find that the Magellanic Clouds are discussed as star clouds in our own galaxy rather than as external galaxies which are satelites of our system. The concluding chapters on the Galactic System and Beyond the Milky Way are excellently treated. One might desire a more detailed discussion of the age of the universe which indicates that we are relatively young (10^9 or 10^{10} years old) and not a decrepit 10^{12} years as Professor Duncan states.

The appendix to the book includes many useful tables which in the earlier editions had been scattered through the book. A table for obtaining the Julian Day for any date from 1940–1990 is an especially welcome addition.

Teachers will wish that the paragraphs were numbered for ease in making assignments, but this is more than compensated for by a large selection of excellent problems involving simple mathematical manipulation.

The book is excellently written, is thoroughly up to date, and includes a nice balance of historical background. It is by far the most attractive general book on astronomy ever written, and the illustrations, many of which were made by Dr. Duncan, are magnificent. The book will undoubtedly be even more popular than the widely used earlier editions, and rightly so.

Carl K. Seyfert
Vanderbilt University

New Booklets

The following recent publications of the National Bureau of Standards are of interest to physicists: RP 1790, *Infrared Emission Spectra of Krypton and Argon*, Curtis J. Humphreys and Earle K. Plyler, 10 cents; RP 1793, *Changes Caused in the Refractivity and Density of Glass by Annealing*, Arthur Q. Tool, Leroy W. Tilton, and James B. Saunders, 10 cents.

Beginning with the October issue, the quarterly journal *Mathematical Tables and Other Aids to Computation* will

publish a new feature section, "Automatic Computing Machinery," designed to disseminate information and news on research and development in the field of high speed automatic calculating machinery. Contributions to this section are invited and should be addressed to Dr. E. W. Cannon, Head of the Mathematics Group, Machine Development Laboratory, National Bureau of Standards, Washington, D. C. Calendar year subscription to the journal is $4; single numbers cost $1.25. Payments should be made to the National Academy of Sciences.

Tracerlog for August 1947, published by Tracerlab, Inc., 55 Oliver Street, Boston 10, Massachusetts, describes the film badge service offered by Tracerlab for health monitoring purposes where employees are exposed to radiation. 16 pages.

Sound Apparatus Company, 233 Broadway, New York 7, New York, recently brought out two bulletins of interest to physicists: *Impedance Vectorgraph* (4 pages) and *Portable Reverberation Analyzer* (12 pages). Because of the highly specialized nature of the latter brochure, the Company is interested in sending it only to people actually engaged in acoustical measuring work.

North American Philips Company, 100 East 42 Street, New York 17, New York, has announced two publications that are available without charge: an 8-page booklet on quality control, and a 4-page folder (R1050) entitled *Geiger-Mueller Counters Applied to Mining.*

The Frontier for June 1947, published by Armour Research Foundation of Illinois Institute of Technology, contains the following featured articles: "The Application of Industrial Research," "Noise and Vibration Reduction," "The Mechanism of Cavitation Erosion," "Crystal Structure of Nicotinamide," and "Contact Angle Measurements in Lithographic Research."

Here and There

Personnel

The University of Wisconsin has announced the retirement in June of Florence E. Allen, Assistant Professor of Mathematics, and John R. Roebuck, Professor of Physics.

Two promotions in the Department of Physics at New York University were announced recently. Morton Hamermesh was elevated to an Associate Professorship and Serge A. Korff to a full Professorship.

Donald A. Quarles, Director of Apparatus Development of Bell Telephone Laboratories, was elected a Vice President at a meeting of the board of directors in July.

Kinichi Watanabe, formerly of the University of Hawaii, has been appointed Assistant Professor in the Department of Physics at Wabash College.

College of Horology and Instrument Technology

A National College of Horology and Instrument Technology is being set up in the Northampton Polytechnic, London, England, to open in October of this year. Full information may be obtained by addressing Mr. R. A. Fell, Northampton Polytechnic, St. John Street, London, E.C. 1.

Institute of Numerical Analysis Established

Plans have been completed for the establishment of one of the newest units of the National Bureau of Standards—the Institute of Numerical Analysis—at the University of California. One of the giant high speed electronic computing machines now under development by the Bureau of Standards will be installed at the Institute when completed. The Institute has two primary functions. The first is research in applied mathematics aimed at developing methods of analysis which will extend the use of the high speed electronic computers. The second is to act as a service group for western industries, research institutions, and government agencies. The service function will include not only the use of the machines for problem solving but also assistance in the formulation of problems in applied mathematics of the more complex and novel types.

Induction Heating Corporation

The Induction Heating Corporation, manufacturer of induction and dielectric heating equipment, has moved to new quarters at 181 Wythe Avenue, Brooklyn 11, New York.

Fellowship on Waterproofed Paper

A comprehensive investigation of waterproofed papers, used for such purposes as packaging, curing of concrete, and moisture barriers in buildings, is being undertaken at the National Bureau of Standards through a fellowship of the Waterproof Paper Manufacturers Association. A long-range research program, extending over at least three years, has been jointly agreed upon. The services of Robert C. Reichel as Research Associate are being provided by the Association, with laboratory facilities and technical assistance furnished by the Bureau.

Applied Mathematics Symposium

The first annual symposium in applied mathematics of the American Mathematical Society was held at Brown University August 2–4. The subject of the symposium was "Non-Linear Problems in Mechanics of Continua." The American Institute of Physics, the American Society of Mechanical Engineers, and the Institute of the Aeronautical Sciences were co-sponsors.

Contractor for Brookhaven Atomic Pile

The H. K. Ferguson Company will build the atomic pile and associated facilities at Brookhaven National Laboratory for atomic research on Long Island, New York. Construction has begun and will be completed in 1948.

Honorary Degrees

The following scientists received honorary degrees recently: John H. Van Vleck, Harvard University, the D.Sc. from the University of Wisconsin; Roger Adams, University of Illinois, the D.Sc. from the University of Pennsylvania; Vannevar Bush, Carnegie Institution of Washington, the D.Sc. from Columbia University.

New Course at Yale

Yale University has announced the inauguration of a new course to be offered this year on the Development of the Sciences. It is designed to aid in "the closer integration of the sciences and the humanities in the college curriculum and to emphasize the contribution of the sciences to liberal education." The course will cover two terms. Representing the science of physics is Henry Margenau.

Pittsburgh Spectroscopy Conference

The Eighth Pittsburgh Conference on Applied Spectroscopy, sponsored by the Spectroscopy Society and the University of Pittsburgh, will be held Thursday to Saturday, November 13 to 15, 1947, inclusive, at the Mellon Institute Auditorium. As in the past, sessions on both emission and absorption spectroscopy are planned. Address inquiries to Mary E. Warga, University of Pittsburgh, Pittsburgh 13, Pennsylvania.

Ionosphere Studies Truck Station

A major problem in the study of long distance radio-wave propagation is the role of the ionosphere in returning radio waves of suitable frequency back to the earth's surface. This study may best be accomplished by pulse techniques as used in radar. In fact, the radar techniques developed during the war were based, in great part, on work of this nature previously developed in connection with ionosphere problems.

Up to the present time most ionosphere studies have been confined to the vertical incidence case for reasons of simplified instrumentation and theoretical analysis. The oblique incidence and practical case has been studied by remote station field intensity recording and by theoretical analysis. But little oblique incidence pulse work has been done up to the present time because of the relatively complex instrumentation required. This is the purpose of the equipment set up in the truck station illustrated on the front cover.

The truck contains complete facilities for recording manually or photographically the pulse emissions from a high power ionosphere transmitter located at the Pennsylvania State College. Moderately high power communications equipment is also included for communication with the college station over distances of several hundred miles. A power unit for supplying the equipment is contained in the trailer at the rear of the truck. The pulse equipment receiving antenna is supported by two sectionalized masts.

For operation, the truck is driven to the place where observations are to be made, the two antennas are erected and, by pushing a button, the power unit is turned on. The equipment is then ready for operation. Experience has shown that this may all be done in about five minutes. Hence, if desired, observations may be conducted at many distances or locations with respect to the transmitting position in a single day.

Several studies utilizing this equipment are currently in progress. This work is being done by Professor A. H. Waynick of the Electrical Engineering Department under contract between the Pennsylvania State College and the Watson Laboratories, U. S. Army Air Forces.

Calendar of Meetings

October

15–18	American Society of Civil Engineers, New Orleans, Louisiana
20–24	American Society for Metals, Chicago, Illinois
21–23	American Standards Association, New York, New York
21–25	American Chemical Society, California Section, San Francisco, California
23–25	Optical Society of America, Cincinnati, Ohio
30–31	National Research Corporation and the American Chemical Society (High Vacuum Symposium), Cambridge, Massachusetts

November

3–5	National Electronics Conference, Chicago, Illinois
3–7	American Institute of Electrical Engineers, Chicago, Illinois
6–7	Society of Automotive Engineers (Fuels and Lubricants Meeting), Tulsa, Oklahoma
7–8	Conference on X-Ray and Electron Diffraction, Mellon Institute of Industrial Research, Pittsburgh, Pennsylvania
9–12	American Institute of Chemical Engineers, Detroit, Michigan
13–15	Conference on Applied Spectroscopy, Mellon Institute, Pittsburgh, Pennsylvania
17–19	Institute of Radio Engineers and the Radio Manufacturers' Association, Engineering Department, Rochester, New York
28–29	American Physical Society, Houston, Texas

December

1–3	Society of Automotive Engineers (Air Transport Engineering Meeting), Kansas City, Missouri
1–5	American Society of Mechanical Engineers, New York, New York
26–31	American Association for the Advancement of Science, Chicago, Illinois
29–31	American Physical Society, Chicago, Illinois